Copernicus
Europe's eyes on Earth

ECMWF

Copernicus Atmosphere Monitoring Service

Validation report of the CAMS near-real time global atmospheric composition service

September – November 2019

Issued by: KNMI
Date: 26 March 2020
Ref: CAMS84_2018SC2_D1.1.1_SON2019

European Commission

Copernicus
Europe's eyes on Earth

IMPLEMENTED BY
ECMWF

Validation report of the CAMS near-real-time global atmospheric composition service. Period September - November 2019

EDITORS:
A. Wagner (MPG), M. Schulz (MetNo), Y. Christophe (BIRA-IASB),
M. Ramonet (LSCE), H.J. Eskes (KNMI)

AUTHORS:
S. Basart (BSC), A. Benedictow (MetNo), Y. Bennouna (CNRS-LA),
A.-M. Blechschmidt (IUP-UB), S. Chabrillat (BIRA-IASB), E. Cuevas (AEMET),
A. El-Yazidi (LSCE), H. Flentje (DWD), K. M. Hansen (AU), U. Im (AU),
J. Kapsomenakis (AA), B. Langerock (BIRA-IASB), A. Richter (IUP-UB),
N. Sudarchikova (MPG), V. Thouret (CNRS-LA), T. Warneke (UBC),
C. Zerefos (AA)

REPORT OF THE COPERNICUS ATMOSPHERE MONITORING SERVICE, VALIDATION SUBPROJECT.

Executive Summary

The Copernicus Atmosphere Monitoring Service (http://atmosphere.copernicus.eu, CAMS) is a component of the European Earth Observation programme Copernicus. The CAMS global near-real time (NRT) service provides daily analyses and forecasts of reactive trace gases, greenhouse gases and aerosol concentrations. This document presents the validation statistics and system evolution of the CAMS NRT service for the period up to 1 December 2019, with a focus on September-November 2019 (SON-2019). Updates of this document appear every 3 months, e.g. Ramonet et al. (2019). A detailed description of the measurement datasets used is provided in Eskes et al. (2019).

This summary is split according to service themes as introduced on the CAMS website: air quality & atmospheric composition, climate forcing, ozone layer and UV. Specific attention is given to the ability of the CAMS system to capture recent events. We focus on the 'o-suite' composition fields, which are the daily analyses and forecasts produced by the IFS (Integrated Forecast System) modelling system at ECMWF, using the available meteorological and atmospheric composition observations which are ingested in the ECMWF 4D-Var assimilation system. The model and assimilation configurations are summarised in section 2. We furthermore assess the impact of the composition observations by comparing the validation results from the 'o-suite' to a 'control' configuration without atmospheric composition data assimilation. Also, the pre-operational delayed-mode analyses and high-resolution forecasts of CO_2 and CH_4 are assessed in this report.

On 9 July 2019, a major upgrade of the CAMS system to version 46R1 took place. Among other things this involved a change from 60 vertical levels to 137 vertical levels. The upgrade is described in more detail in section 2, and special attention is given in this validation report to changes in performance linked to this upgrade.

The o-suite data delivery for the period SON-2019 was very good, with an on-time percentage of 98.9%.

Air quality and atmospheric composition

Tropospheric ozone (O₃)

CAMS o-suite ozone is validated with surface and free tropospheric ozone observations from the GAW and ESRL networks, IAGOS airborne data, ozone sondes and IASI tropospheric ozone retrievals. For free tropospheric ozone against ozone sondes the o-suite modified normalized mean biases (MNMBs) are on average small, ±10% over the Northern Hemisphere (NH), between ±30% for stations in the Tropics, and ±20% for the Arctic in more recent years (Fig. S.1). Over Antarctica o-suite biases are observed between 0% and +30% for recent years, whereas the control run shows negative biases. For SON 2019 good agreement is found over the NH mid latitudes in the free troposphere.

For the IAGOS validation, the biases remain stable since 16 August, when the replacement of the instrument presenting issues in ozone measurements took place. Over Frankfurt, a good agreement with IAGOS is found for surface to free troposphere ozone, with mostly only small overestimations
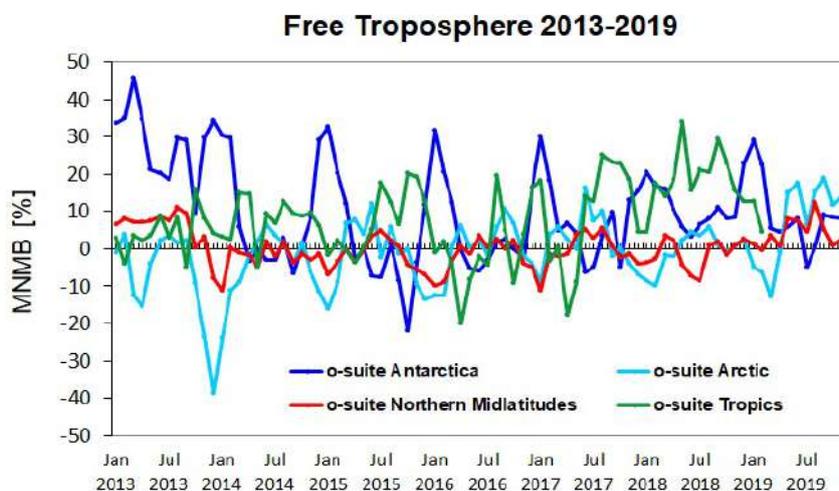
Figure S.1: Time series of MNMB of ozone in the o-suite, compared against ozone sondes, averaged over different latitude bands. The free troposphere is defined here as the layer between 750 and 300 hPa.

in the lowest layers from both the o-suite and control run and the smallest bias in the free troposphere. In the UTLS region, the biases are larger and the results from the models differ, ozone is mostly overestimated by the o-suite whereas the control run often underestimates. Good agreement with the IAGOS free troposphere ozone profiles is also found over the other regions of the world, with sometimes large differences in the UTLS region.

The validation with IASI satellite data shows that the o-suite run captures both, high and low O_3 values relatively well and is in good agreement with the observations, showing MNMBs within 5%. The control run is mainly positively biased (up to 20% over the high northern latitudes). Slightly negative biases appear over the biomass burning area in Africa (within 10%).

In comparison with surface observations we find a steady improvement of the o-suite over the past 5 years over European GAW stations. Biases are within 20% (the Arctic is discussed below). The o-suite has biases mostly smaller than 10% for surface ozone for Europe during September to November 2019. Both runs overestimate O_3 minimum observations for Asia with MNMBs up to 20% for the o-suite and the control run. For the tropics, surface ozone is overestimated with MNMBs within 20%. Both runs can reproduce Antarctic surface observations with MNMBs within ±20%.

Tropospheric Nitrogen dioxide (NO$_2$)

Model validation, with respect to SCIAMACHY/Envisat NO_2 data before April 2012 and GOME-2/MetOp-A NO_2 data afterwards, shows that tropospheric NO_2 columns are well reproduced by the NRT model runs, indicating that emission patterns and NO_x photochemistry are generally well represented, although modelled shipping signals are more pronounced than in the satellite retrievals. Tropospheric NO_2 columns over some local emission hotspots (e.g. Moscow, and Red Basin in China) are overestimated, while wintertime and springtime values over Europe around Benelux are underestimated. Since December 2014, the agreement between satellite retrievals and model results for time series over East-Asia and Europe is better than for previous years (Fig. S.2), as
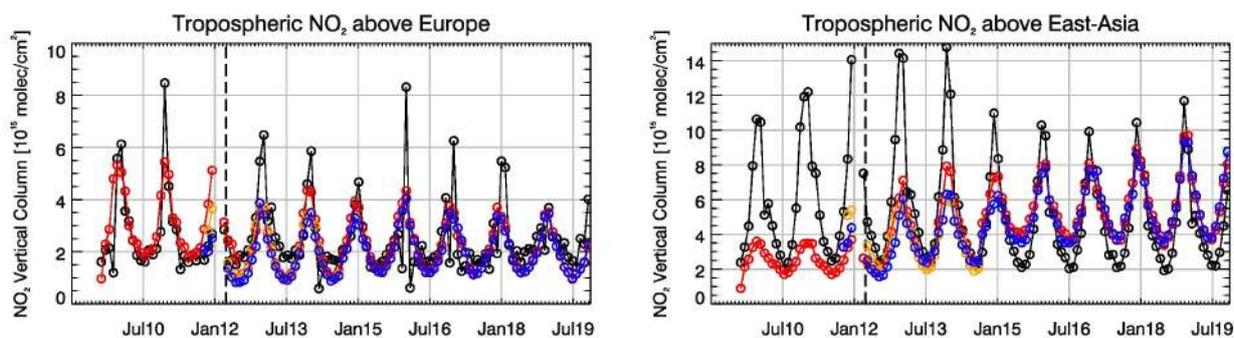
Figure S.2: Time series of tropospheric NO$_2$ columns from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for Europe and East-Asia. Black lines show the observations, red shows the o-suite, blue lines show CAMS control results including older configurations from the MACC projects before September 2014, orange shows the forecast from the MACC project, based on the Mozart model.

observed columns of NO$_2$ decreased in 2014, likely associated with reduced emissions, and (in contrast to the observations) simulated values show an increase over the whole time-series available. Between spring and autumn, the models regularly show an overestimation over several regions with boreal forest fire activity (Canada, Alaska, Siberia).

Tropospheric Carbon Monoxide (CO)

Model validation with respect to GAW network surface observations, IAGOS airborne data, FTIR observations (NDACC and TCCON) and MOPITT / IASI satellite retrievals reveals that the absolute values, latitude dependence and seasonality, as well as day-to-day variability of CO can be reproduced well by the CAMS-global analyses and forecasts. Biases are between 4% and -15% for European GAW stations, and up to -25% in Asia.

The comparison with NDACC data shows, that the model upgrade (60 to 137 levels) implemented in July 2019 changes the overall biases in both the troposphere and stratosphere. The bias for the tropospheric columns becomes -7% in SON (-4% in JJA) and is larger than the reported measurement uncertainty. The stratospheric column bias reduces to +2% in SON (+6% in JJA) and now falls within the measurement's uncertainty.

The comparisons show that the last model update improved the comparison between the o-suite AN and o-suite 1d FC models with TCCON data, while the control model shows a low bias. Also, the variation of the CO after the model update is captured well by the o-suite analysis and o-suite 1-day forecast, as shown exemplary for the TCCON site Orleans. The available data for the reporting period is limited to the Northern Hemisphere, except for the site Reunion.

According to IAGOS observations, CO is mostly underestimated over Frankfurt by both the o-suite and control run and the largest bias is found in the lowest layers. The performance of the two runs is similar in the lowest layers, while for the free troposphere the performance of the o-suite is slightly better than that of the control run. For most other regions of the world, the results are similar to those of Europe.
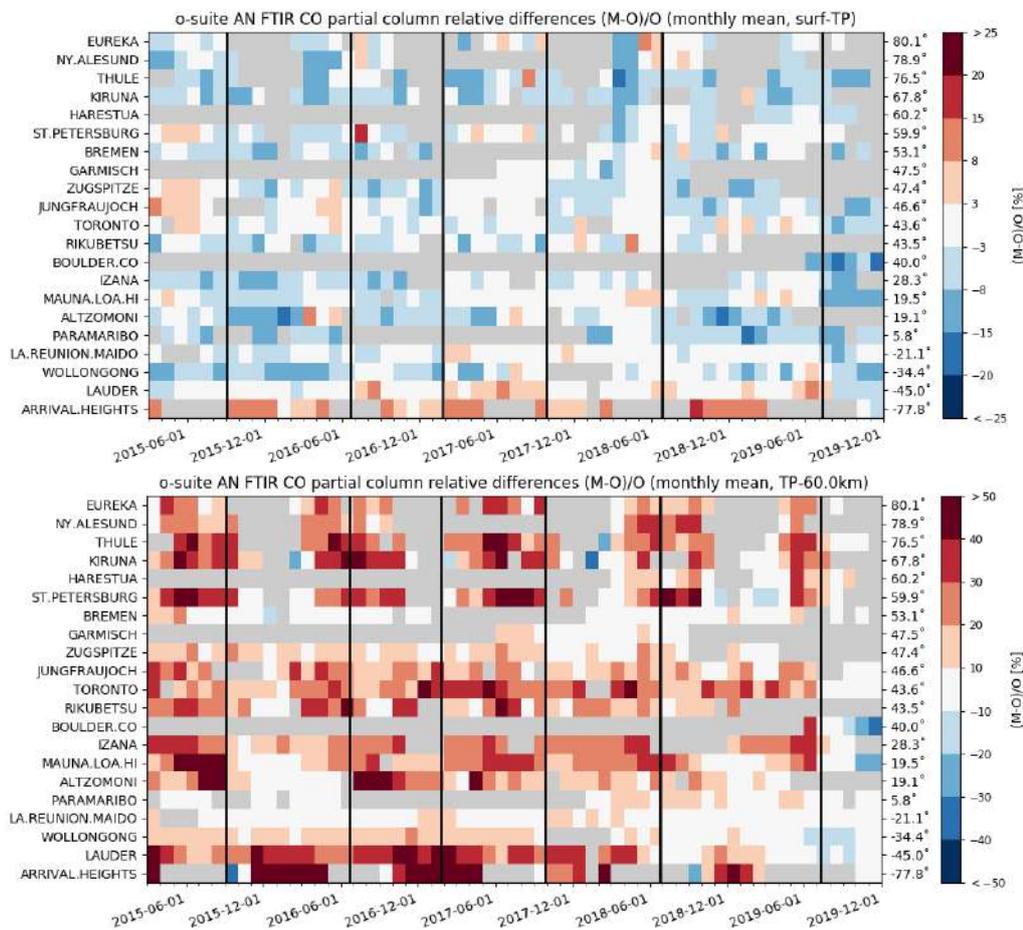
Figure S.3: Monthly relative mean bias from the NDACC FTIR comparisons for tropospheric (top) and stratospheric (bottom) CO columns (%) for the considered period up to 1 December 2019. Model upgrades are indicated in black vertical lines. The overall uncertainty for the CO measurements is approximately 3% on the tropospheric columns and 10% for the stratospheric columns. The o-suite analysis averaged bias in tropospheric columns increased to -7% for SON 2019. The bias in the stratosphere reduced to +2% and lies within the measurement's uncertainty. Stations are sorted with decreasing latitude (northern to southern hemisphere).

The comparisons with MOPITT and IASI confirm these findings. The o-suite is generally underestimating the satellite data by about 10% with some regional exceptions where the negative bias reaches 20% (mostly over the land). After the update from 60 to 137 levels in the second half of the year 2019, we see the following changes:

- Enhanced negative biases over the US and Europe in the o-suite run and more pronounced underestimations in the control run.
- Improvement of the o-suite results over East and South Asia (bias is almost zero).
- General change of bias sign in the control run from positive to negative over the Siberian fire region and enhanced negative biases in the o-suite run.
- Strong increase of negative bias for the control run over the Alaskan region.
- Stronger underestimation over South Africa in both, the o-suite and control run and stronger underestimation over North Africa for the o-suite.
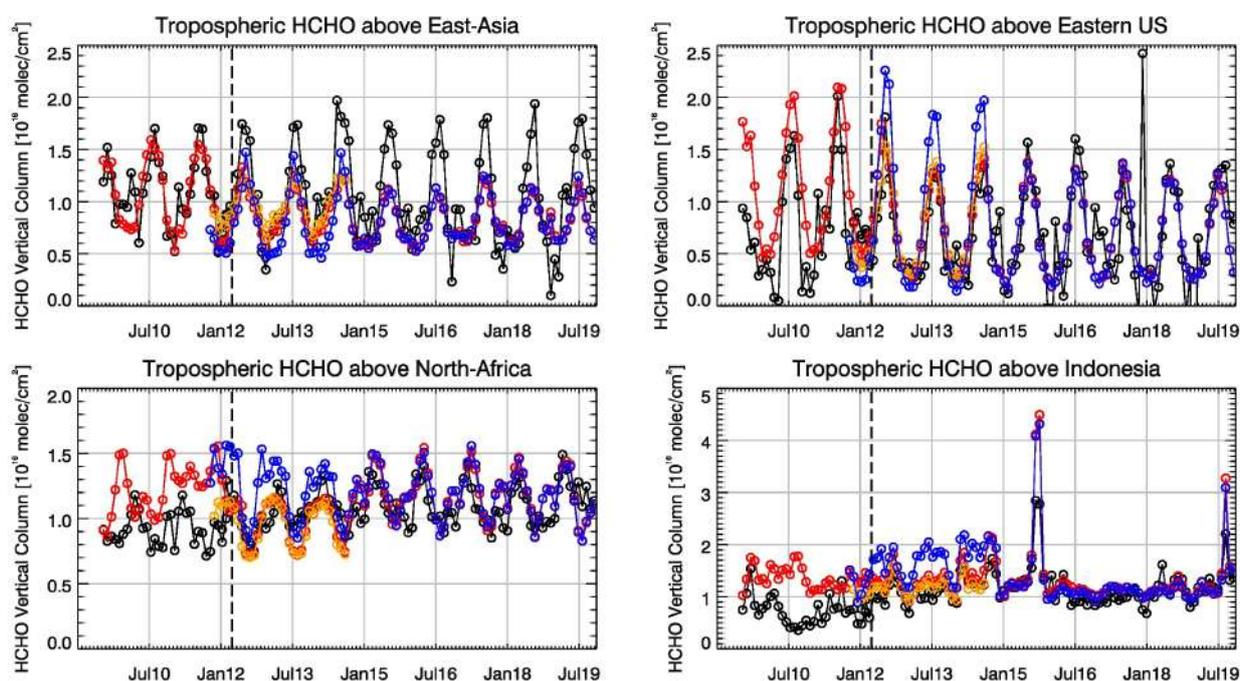
Figure S.4: Time series of average tropospheric HCHO columns [10^{16} molec cm^{-2}] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012 onwards) compared to model results for different regions. Black lines show the observations, red shows the o-suite, blue lines show CAMS control results including older configurations from the MACC projects before September 2014, orange shows MACC_fcnrt_MOZ from the MACC projects. The regions are: East-Asia (25-40°N, 110-125°E), Eastern US (30-40°N, 75-90°W), Northern Africa (0-15°N, 15°W-25°E) and Indonesia (5°S-5°N, 100-120°E). Vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012

Formaldehyde

Model validation, with respect to SCIAMACHY/Envisat HCHO data before April 2012 and GOME-2/MetOp-A HCHO data afterwards (Fig. S.4), shows that modelled monthly HCHO columns represent well the magnitude of oceanic and continental background values and the overall spatial distribution in comparison with mean satellite HCHO columns. Compared to GOME-2 satellite retrievals, an overestimation of values regularly occurs over Australia and Central Africa, which could be both related to biogenic emissions and fire emissions. For time series over East-Asia and the Eastern US, both regions where HCHO columns are probably dominated by biogenic emissions, models and retrievals agree rather well, but the yearly cycle over East-Asia is underestimated by the models.

Aerosol

We estimate that the o-suite aerosol optical depth showed an average positive bias in the latest three months of +16%, measured as modified normalized mean bias against daily Aeronet (V3 level 1.5) sun photometer data. The 3-day forecasted aerosol distribution shows 14% less aerosol optical depth (AOD) than that from the initial forecast day, as shown in Fig. S.5-a. Spatiotemporal correlation, shown in Fig. S.5-b, shows month-to-month variation in SON 2019 similar to autumn 2018, indicating the simulation reproduces approximately 68% of the day to day AOD variability
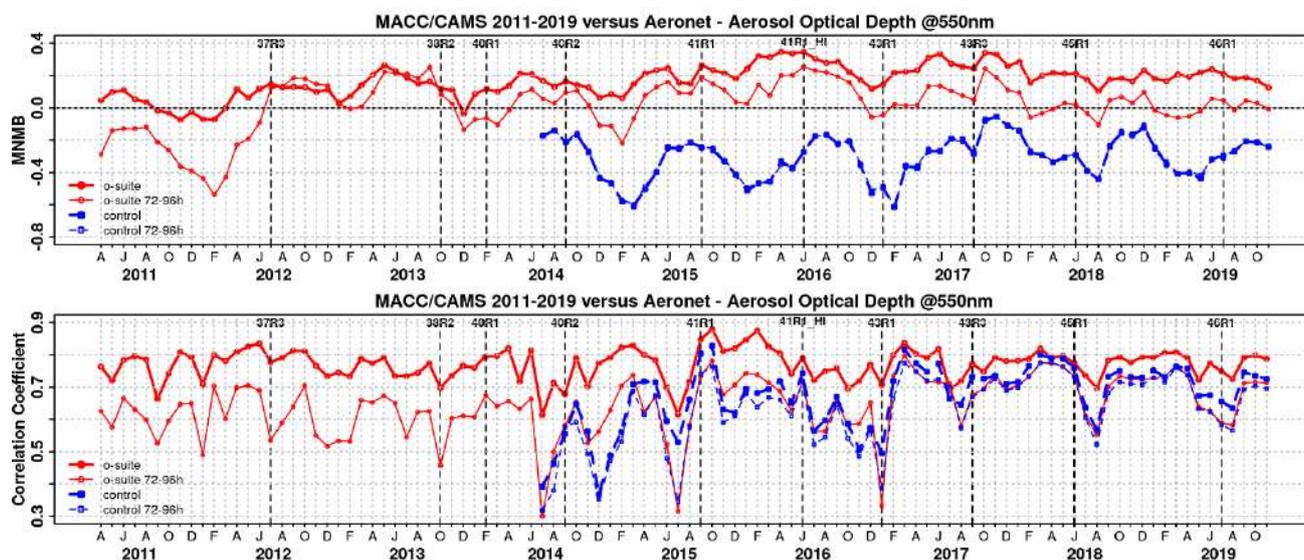
Figure S.5. Aerosol optical depth at 550nm in IFS 00Z model simulations for April 2011 – November 2019 against daily matching Aeronet Version3 level 1.5 data. a) Modified normalized mean bias (MNMB); o-suite (thick red curve); o-suite at last forecast day (light red curve); Control (blue dashed); Control at last forecast day (light blue dashed); b) Corresponding correlation coefficient. Model version changes are marked as vertical bars.

across all Aeronet stations. The o-suite forecast at +3 days shows slightly lower correlation, as a consequence of imperfect forecasted meteorology and fading impact of the initial assimilation of MODIS AOD and MODIS fire info on model performance. The o-suite forecast running each day at 12UTC shows almost identical performance as the forecast starting at 00UTC.

The AOD performance of the o-suite with respect to the AERONET data exhibits no pronounced seasonal cycle but somewhat less correlation in late summer. Since October 2017, the largest contributions to global AOD come from organics and sea salt. Sea salt AOD increased further due to the penultimate model upgrade in June 2018 with the new sea salt emission scheme activated, while dust AOD became lower compared to earlier years. With the coupling of chemistry and aerosol schemes for sulphur in the latest upgrade in July 2019, there is an increase of SO4 especially in the northern hemisphere.

The aerosol Ångström exponent (AE) contains information about the size distribution of the aerosol, and implicitly about composition. The o-suite AE became more positive indicating a change to slightly more fine particles since the model upgrade to version 45R1 in June 2018.

PM10 and PM25, as defined by the IFS aerosol model, are evaluated against an average from rural and background site data in the period 2000-2009 at 160 sites in North America and Europe. This indicates that PM10 concentrations exhibit on average in the latest period an underestimation with MNMB bias of -35% in Europe and of -7% in North America. PM25 concentrations are underestimated with -35% in Europe and overestimated with 28% in North America. Consistent with this finding a higher positive bias is also found for AOD in North America than in Europe. The fraction of PM simulated data within a factor 2 of observed values stayed similar since September 2017 for both PM10 and PM25. PM25 seems to have deteriorated compared to periods before mid
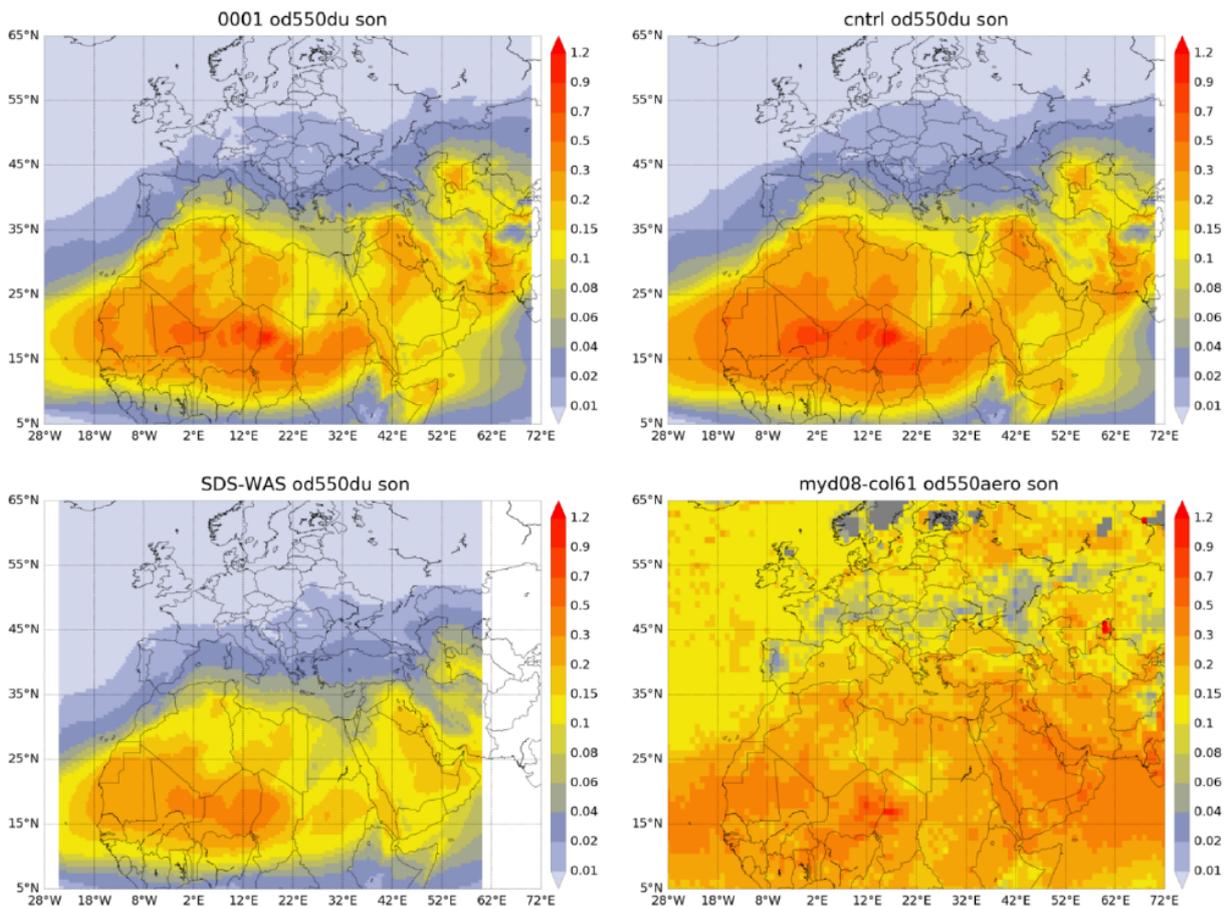
Figure S.6: Averaged DOD 24h forecast from o-suite (top left) and control (top right), DOD of the multi-model SDS-WAS Median product (bottom left) as well as AOD from MODIS/Aqua Collection 6.1 Level 3 combined Dark Target and Deep Blue product (bottom right) for the study period.

2017, while PM10 shows an improvement. However, with the latest model version upgrade in July 2019, the PM25 has improved significantly during September-November 2019. Over Northern Africa, the Middle East and Europe, the dust component of the CAMS o-suite shows lower season values than the control run, which is in general higher than the SDS-WAS multi-median product. The CAMS o-suite does reproduce the dust transport over the North Atlantic region. However, the maximum dust activity is shifted to Mauritania, Mali, Niger and Algeria border and Sudan in addition to Chad, as shown in MODIS. Also, DOD over Iraq and in the Mediterranean basin appears overestimated in comparison with the SDS-WAS multi-model ensemble. These changes in dust activity in the main source regions are linked to the new dust module implemented in the operational CAMS model since early-July 2019.

From September to November, the o-suite and control experiments reproduce the daily variability of AERONET direct-sun observations with a correlation coefficient of 0.77, averaged over all the AERONET sites; this is lower than in the case of the SDS-WAS multi-model product shows a correlation coefficient of 0.82. Regarding mean bias (MB), the o-suite tends to reduce the overestimations observed in control with a MB of -0.01 for o-suite and 0.03 for control run, meanwhile the SDS-WAS multi-model underestimates (MB of -0.03) this AERONET dust-filtered observations.

The comparison of 1 to 3-day forecasts shows that the prediction is stable during the 3-days forecasts in comparison with AERONET direct-sun observations with correlation coefficients of 0.77 (0.77), 0.75 (0.76), and 0.75 (0.75) respectively for 24, 48 and 72h forecasts for all the sites for o-suite (control).

System performance in the Arctic

The CAMS model runs are validated using surface ozone measurements from the ESRL-GMD and the IASOA networks and ozone concentrations in the free troposphere and UTLS are evaluated using balloon sonde measurement data.

From September to November 2019 the simulations of the surface ozone concentrations are on average in good agreement with the observations with MNMB between -11% and 15%.

During September – November 2019 there is an overestimation of ozone concentrations in the Arctic free troposphere for the o-suite (MNMB = 15% – 19%) and for the control run (MNMB = 3% – 8%) as well as in the UTLS (MNMB up to 30% for the o-suite). The larger positive biases might be due to the stop in assimilation of OMPS data From May 2019 onwards.

Total column O_3 is in good agreement with the observations obtained from IASI, showing MNMBs within 5%. The IASI sensitivity is the lowest over the cold surfaces such as the Greenland ice sheet where IASI O_3 values are positively biased by up to 20%.

Comparison with FTIR observations from the NDACC network shows that the CO tropospheric columns are underestimated at Thule with bias -11%, with larger negative bias for the control run (-22%), while the stratospheric column is in better agreement with bias of 0.5% for the o-suite and -8% for the control run.

Comparison with MOPITT versions 7 shows that modelled CO total columns are in relatively good agreement with the satellite retrievals with low bias in the Arctic (±10%).

System performance in the Mediterranean

The CAMS o-suite reproduces the daily variability of AERONET direct-sun observations. In the Western, Central and Eastern Mediterranean, the correlation coefficient of the o-suite (0.78, 0.71 and 0.68) is improved by the assimilation compared to the control run (0.70, 0.64 and 0.64) during autumn. CAMS overestimated the AERONET observations in the Mediterranean Basin in control (MB of 0, 0.01 and -0.01 for Western, Central and Eastern Mediterranean regions respectively) and o-suite (MB of 0.02, 0.03 and 0.01). The highest peaks on CAMS AOD simulations are linked to desert dust intrusions occurring during the whole season in the entire Mediterranean basin. In the second half of October, high AOD values are observed in SEDE BOKER (up to 0.5) that are not associated with natural aerosols (i.e. dust or sea-salt). This can be associated with the latest model upgrade with the improved description of the sulphur cycle.

The PM10 and PM2.5 results of the CAMS o-suite and control run show similar skill scores in comparison with EIONET-Airbase observations. The CAMS o-suite tends to underestimate the PM10 and PM2.5 EIONET-Airbase observations. The o-suite shows higher overestimations in PM10 and PM2.5 (with a MB of -2.93 and -1.64 μg/m³, respectively) than control (with a MB of -4.43 and -2.68 μg/m³, respectively). Overestimations are observed in Southern European sites, in particular in Italy

for the PM10 fraction. The upgrade of the CAMS model during July 2019 led to an increase of the coarse particles at surface levels.

The model is compared to surface O_3 observations from the AirBase network. Our analysis shows that model MNMBs vary between -14% and 30% depending on the station. Temporal correlation coefficients between simulated and observed surface ozone for both the o-suite and control runs are highly significant over the entire Mediterranean from Gibraltar to Cyprus.

Climate forcing

Greenhouse gases

CO2 and CH4 surface concentrations from ICOS network, and total or partial columns from TCCON and NDACC stations have been used to validate the analysis and high-resolution forecast experiments.

According to ICOS stations, the bias on CH_4 surface measurements is clearly dependent on latitude, at least in Europe. We observe a positive bias (10 to 50 ppb) at high latitude, and a negative one in Southern Europe (-10 to -40 ppb). Column measurements (TCCON and NDACC) indicate negative biases (except for urban sites close to Los Angeles), but also appear to be dependent on altitude since NDACC measurements in the stratosphere show a slight overestimation compared to the measurement uncertainty.

The surface and total column measurements indicate an overestimation of the amplitude of the CO2 seasonal cycle in the northern hemisphere by ±1%. Both surface and total column measurements indicate a maximum bias between July and October 2019 of up to +10 ppm.

Ozone layer and UV

Ozone partial columns and vertical profiles

Ozone columns and profiles have been compared with the following observations: vertical profiles from balloon-borne ozonesondes; ground-based remote-sensing observations from the NDACC (Network for the Detection of Atmospheric Composition Change, http://www.ndacc.org); and satellite observations by 3 instruments (OMPS-LP, ACE-FTS and SAGE-III). Furthermore, the o-suite analyses are compared with those delivered by the independent assimilation system BASCOE.

Compared to ozone sondes (Fig. S.7) the o-suite stratospheric O_3 partial pressures are slightly overestimated in all latitude bands (MNMB between -1 and +14%) except above the Antarctic.

Comparisons with the NDACC network include 19 stations for FTIR, 16 stations for UVVIS stratospheric columns, microwave profiles for Ny Alesund (78.9°N) and Bern (47°N) and LIDAR profiles at Hohenpeissenberg (47.8°N) and Observatoire Haute Provence (OHP), France (43°N) and Mauna Loa, Hawai (19.5°N). The comparisons show a general good agreement with the o-suite, with small performance differences between AN and 1d forecasts. At the tropical sites the 1d FC performs significantly worse since the June 2016 update of the o-suite. This is confirmed by FTIR and the Mauna Loa LIDAR measurements.
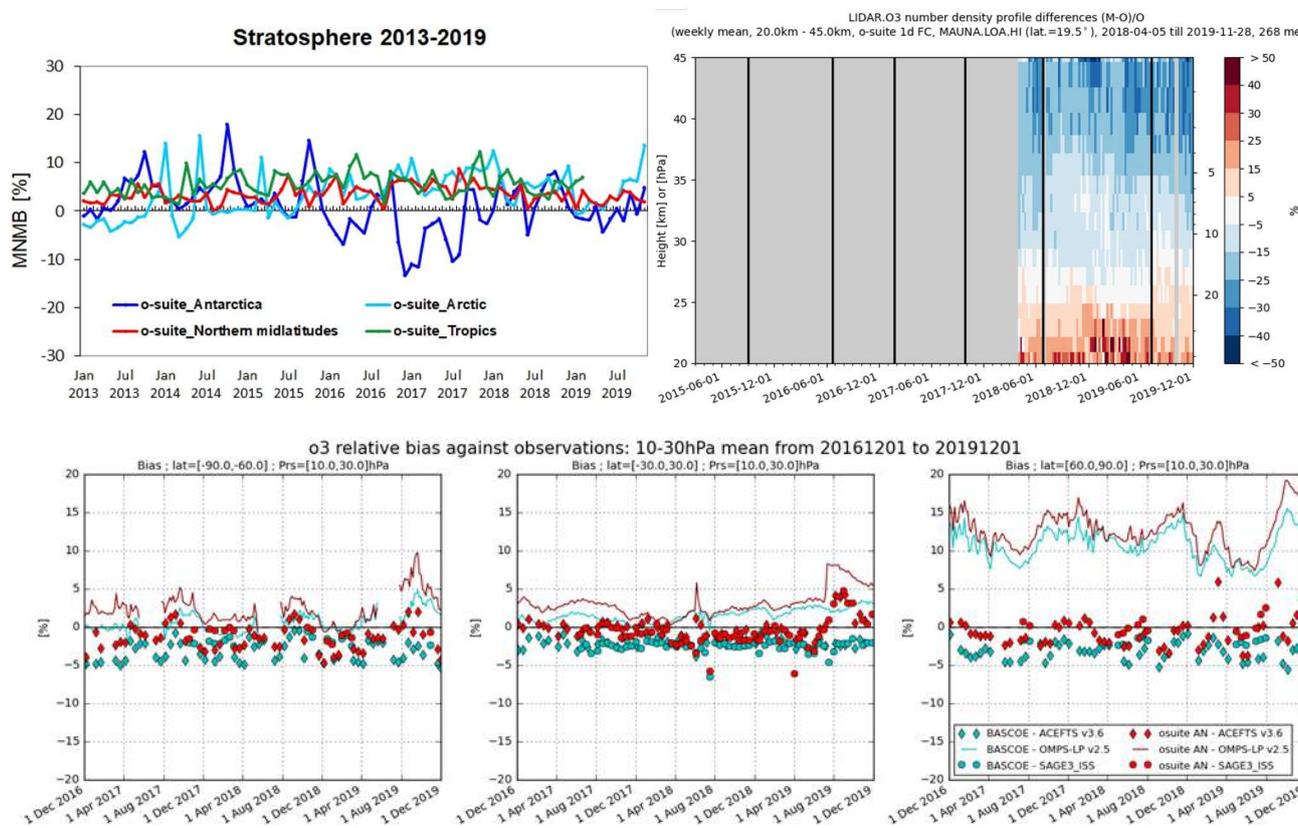
Figure S.7. Top-left: MNMBs (%) of ozone in the stratosphere from the o-suite against aggregated ozonesonde data in the Arctic (light blue), Antarctic (dark blue) northern midlatitudes (red) and tropics (green) from 2013 to November 2019. The stratosphere is defined as the altitude region between 60 and 10 hPa in the tropics and between 90 and 10 hPa elsewhere. Top-right: Comparison of the weekly mean profile bias between the O_3 mixing ratios of the 1-d forecast and the NDACC LIDAR at Mauna Loa. Bottom: Time series comparing models to observations for the period 2016-12-01 to 2019-12-01 in the upper stratosphere (10-30hPa averages): o-suite analyses (red) and BASCOE (cyan) vs OMPS-LP (solid), ACE-FTS (diamonds) and SAGE-III (bullets) for the Antarctic (left), the Tropics (centre) and the Arctic (right). Shown is the normalized mean bias (model-observation)/observation (%).

The comparison with independent satellite observations is generally in good agreement for the considered period: for ACE-FTS, the NMB is mainly within 10% between 5km and 40km, and mostly within 5% between 15km and 35km except in the tropics; for SAGE-III, the NMB is mainly within 10% between 15km and 40km. OMPS-LP has less regular profiles, but the NMB still remain within 15% for most parts of the 20-40 km range. Since the upgrade to cycle 46R1 and the new vertical grid of 137 levels on July 9th, a systematic overestimation of up to 10% around 30km (20hPa) is present in the NMB profiles.

Other stratospheric trace gases

Due to the lack of stratospheric chemistry in the C-IFS-CB05 scheme, the only useful product in the stratosphere is ozone. Other species, like NO_2, have also been evaluated but the results are only indicative.

Events

Heat waves in Europe, June-July 2019: The CAMS NRT performance was validated for the two major heat waves over western and central Europe, during 26-30 June 2019 and 24-27 July 2019 against Airbase observations. In a large part of western and central Europe air temperatures exceeds the JJA 1981-2100 climatological mean up to 8 ºC. In both cases the heat waves lead to high ozone, higher than the JJA 2019 mean by up to 75%, with the maximum exceedance observed over France, Switzerland and Belgium. The CAMS NRT products do reproduce the observed anomaly pattern but underestimate the peak values.

Iberian Peninsula dust event in early-September 2019: In mid-September 2019, the MODIS satellite instrument detected two dust outbreaks that reached the Iberian Peninsula. One outbreak occurred on 13-17 September June, originating from Algeria and transported towards Western Mediterranean, hitting the Balearic Islands and Southern France, and moving to the west crossing the Iberian Peninsula. The CAMS o-suite AOD did timely reproduce the spatial distribution of the two dust plumes as shown the comparison with MODIS/Aqua AOD comparison despite the model tendency to overestimate the observed maximum values. During this event, the o-suite is predicting PM10 values of over 300µg/m3 in large intrusions over the Iberian Peninsula, overestimating the PM10 observations in Spain and Portugal.

Dust event over Germany, 23 April 2019: Due to the prolonged dryness, the vegetation activity in eastern Poland was very weak. Accompanied by wind gusts, dust has been uplifted and was transported along the pressure gradient in the boundary layer (up to 2500m). Shortly after (1 to 1,5 days) Saharan dust may have been transported to the region.

Table of Contents

1. Introduction

The Copernicus Atmosphere Monitoring Service (CAMS, http://atmosphere.copernicus.eu/) is a component of the European Earth Observation programme Copernicus. The CAMS global near-real time (NRT) service provides daily analyses and forecasts of trace gas and aerosol concentrations. The CAMS near-real time services consist of daily analysis and forecasts with the ECMWF IFS system with data assimilation of trace gas concentrations and aerosol properties. This document presents the system evolution and the validation statistics of the CAMS NRT global atmospheric composition analyses and forecasts. The validation methodology and measurement datasets are discussed in Eskes et al. (2015).

In this report the performance of the system is assessed in two ways: both the longer-term mean performance (seasonality) as well as its ability to capture recent events are documented. Table 1.1 provides an overview of the trace gas species and aerosol aspects discussed in this CAMS near-real time validation report. This document is updated every 3 months to report the recent status of the near-real time service. The report covers results for a period of at least one year to document the seasonality of the biases. Sometimes reference is made to other model versions or the reanalysis to highlight aspects of the near-real time products.

This validation report is accompanied by the "Observations characterization and validation methods" report, Eskes et al. (2019), which describes the observations used in the comparisons, and the validation methodology. This report can also be found on the global validation page, http://atmosphere.copernicus.eu/user-support/validation/verification-global-services.

Key CAMS NRT products and their users are: Boundary conditions for regional air quality models (e.g. AQMEII, air quality models not participating in CAMS); Long range transport of air pollution (e.g. LRTAP); Stratospheric ozone column and UV (e.g. WMO, DWD); 3D ozone fields (e.g. SPARC). As outlined in the MACC-II Atmospheric Service Validation Protocol (2013) and MACC O-INT document (2011), relevant user requirements are quick looks of validation scores, and quality flags and uncertainty information along with the actual data. This is further stimulated by QA4EO (Quality Assurance Framework for Earth Observation, http://www.qa4eo.org) who write that "all earth observation data and derived products is associated with it a documented and fully traceable quality indicator (QI)". It is our long-term aim to provide such background information. The user is seen as the driver for any specific quality requirements and should assess if any supplied information, as characterised by its associated QI, are "fit for purpose" (QA4EO task team, 2010).

CAMS data are made available to users as data products (grib or netcdf files) and graphical products from ECMWF, accessible through the catalogue on http://atmosphere.copernicus.eu/.

A summary of the system and its recent changes is given in section 2. Subsequent sections give an overview of the performance of the system for various species, and during recent events. Routine validation results can be found online via regularly updated verification pages, http://atmosphere.copernicus.eu/user-support/validation/verification-global-services.
Table 1.2 lists all specific validation websites that can also be found through this link.

Table 1.1: Overview of the trace gas species and aerosol aspects discussed in this CAMS near-real time validation report. Shown are the datasets assimilated in the CAMS analysis (second column) and the datasets used for validation, as shown in this report (third column). Green colours indicate that substantial data is available to either constrain the species in the analysis, or substantial data is available to assess the quality of the analysis. Yellow boxes indicate that measurements are available, but that the impact on the analysis is not very strong or indirect (second column), or that only certain aspects are validated (third column).

Species, vertical range	Assimilation	Validation
Aerosol, optical properties	MODIS Aqua/Terra AOD PMAp AOD	AOD, Ångström: AERONET, GAW, Skynet, MISR, OMI, lidar, ceilometer
Aerosol mass (PM10, PM2.5)	MODIS Aqua/Terra	European AirBase stations
O_3, stratosphere	MLS, GOME-2, OMI, SBUV-2, TROPOMI	Sonde, lidar, MWR, FTIR, OMPS, ACE-FTS, OSIRIS, BASCOE and MSR analyses
O_3, UT/LS	MLS	IAGOS, ozone sonde
O_3, free troposphere	Indirectly constrained by limb and nadir sounders	IAGOS, ozone sonde, IASI
O_3, PBL / surface		Surface ozone: WMO/GAW, NOAA/ESRL-GMD, AIRBASE
CO, UT/LS	IASI, MOPITT	IAGOS
CO, free troposphere	IASI, MOPITT	IAGOS, MOPITT, IASI, TCCON
CO, PBL / surface	IASI, MOPITT	Surface CO: WMO/GAW, NOAA/ESRL
NO_2, troposphere	OMI, GOME-2, partially constrained due to short lifetime	TROPOMI, SCIAMACHY, GOME-2, MAX-DOAS
HCHO		TROPOMI, GOME-2, MAX-DOAS
SO_2	GOME-2 (Volcanic eruptions)	
Stratosphere, other than O_3		NO_2 column only: SCIAMACHY, GOME-2
CO_2, surface, PBL		ICOS
CO_2, column	GOSAT	TCCON
CH_4, surface, PBL		ICOS
CH_4, column	GOSAT, IASI	TCCON

Table 1.2: Overview of quick-look validation websites of the CAMS system.

Reactive gases – Troposphere

IAGOS tropospheric ozone and carbon monoxide:
    http://www.iagos.fr/cams/
Surface ozone from EMEP (Europe) and NOAA-ESRL (USA):
    http://www.academyofathens.gr/cams
Tropospheric nitrogen dioxide and formaldehyde columns against satellite retrievals:
    http://www.doas-bremen.de/macc/macc_veri_iup_home.html
Tropospheric CO columns against satellite retrievals:
    http://www.mpimet-cams.de
GAW surface ozone and carbon monoxide:
https://atmosphere.copernicus.eu/charts/cams_gaw_ver/v0d_gaw_oper_operfc_nrt_sites?facets=undefined&time=2018060100,0,2018060100&fieldpair=CO&site=cmn644n00

Reactive gases - Stratosphere

Stratospheric composition:
    http://www.copernicus-stratosphere.eu
NDACC evaluation in stratosphere and troposphere (the NORS server)
    http://nors-server.aeronomie.be

Aerosol

Evaluation against Aeronet stations:
    http://aerocom.met.no/cams-aerocom-evaluation/
    More in-depth evaluations are available from the Aerocom website.
WMO Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS)
    model intercomparison and evaluation:
    http://sds-was.aemet.es/forecast-products/models
Aeronet verification of CAMS NRT forecasts:
    https://atmosphere.copernicus.eu/charts/cams_aeronet_ver/?facets=undefined&
    time=2019020100,0,2019020100&site=ARM_Graciosa

Satellite data monitoring
Monitoring of satellite data usage in the Near-Real-Time production: https://atmosphere.copernicus.eu/charts/cams/cams_satmon?facets=undefined&time=2016071800&Parameter=AURA_MLS_profile_Ozone_1_GLOBE

Naming and color-coding conventions in this report follow the scheme as given in Table 1.3.

Table 1.3. Naming and colour conventions as adopted in this report.

Name in figs	experiment	Colour
{obs name}	{obs}	black
o-suite D+0 FC	0001	red
control	gsyg	blue
GHG high-resolution run	gqpe / ghqy	orange
GHG global analysis	gqiq	green

2. System summary and model background information

The specifics of the different CAMS model versions are given below (section 2.1) including an overview of model changes. Other systems used in CAMS are listed in section 2.2. An overview of products derived from this system is given in section 2.3. Timeliness and availability of the CAMS products is given in section 2.4.

2.1 System based on the ECMWF IFS model (the o-suite and control run)

Key model information is given on the CAMS data-assimilation and forecast run o-suite and its control experiment, used to assess the performance of the assimilation. The forecast products are listed in Table 2.1. Table 2.2 provides information on the satellite data used in the o-suite. Further details on the different model runs and their data usage can be found at
http://atmosphere.copernicus.eu/documentation-global-systems.

Table 2.1: Overview of model runs assessed in this validation report.

Forecast system	Exp. ID	Brief description	Upgrades (e-suite ID)	Cycle
o-suite	0001	Operational CAMS DA/FC run	20190709-present 20180626-20190708 20170926-20180625 20170124-20170926 20160621-20170124 20150903-20160620 20140918-20150902	46R1 45R1 43R3 43R1 41R1 41R1 40R2
Control	h7c4 gzhy gsyg gnhb gjjh geuh g4o2	control FC run without DA	20190709-present 20180626-20190708 20170926-20180625 20170124-20170926 20160621-20170124 20150901-20160620 20140701-20150902	46R1 45R1 43R3 43R1 41R1 41R1 40R2
GHG run	ghqy gf39	High resolution T1279, NRT CO_2 and CH_4 without DA	20160301-20170621 20150101-20160229	
	gqpe	High resolution Tco1279 (~9km) NRT CO_2, CH_4 and linCO forecast, initialized from GHG analysis gqiq and CAMS operational CO analysis	20170101-present	
	gqiq	GHG analysis Tco399 (~25km)	20170101-present	

Table 2.2: Satellite retrievals of reactive gases and aerosol optical depth that are actively assimilated in the o-suite.

Instrument	Satellite	Provider	Version	Type	Status
MLS	AURA	NASA	V4	O3 Profiles	20130107 -
OMI	AURA	NASA	V883	O3 Total column	20090901 -
GOME-2A	Metop-A	Eumetsat	GDP 4.8	O3 Total column	20131007 - 20181231
GOME-2B	Metop-B	Eumetsat	GDP 4.8	O3 Total column	20140512 -
SBUV-2	NOAA-19	NOAA	V8	O3 21 layer profiles	20121007 -
OMPS	Suomi-NPP	NOAA / EUMETSAT		O3 Profiles	20170124 - 20190409
TROPOMI	Sentinel-5P	ESA		O3 column	20181204-
IASI	MetOp-A	LATMOS/ULB Eumetsat	-	CO Total column	20090901 - 20180621 20180622 - 20191118
IASI	MetOp-B	LATMOS/ULB Eumetsat	-	CO Total column	20140918 - 20180621 20180622 -
IASI	MetOp-C	Eumetsat		CO total column	20191119 -
MOPITT	TERRA	NCAR	V5-TIR V7-TIR V7-TIR Lance V8-TIR	CO Total column	20130129 - 20160124 - 20180626 20180626 20190702
OMI	AURA	KNMI	DOMINO V2.0	NO2 Tropospheric column	20120705 -
GOME-2A/2B	METOP A/B	Eumetsat	GDP 4.8	NO2 Tropospheric column	20180626 -
OMI	AURA	NASA	v003	SO2 Tropospheric column	20120705-20150901
GOME-2A/2B	METOP A/B	Eumetsat	GDP 4.8	SO2 Tropospheric column	20150902 -
MODIS	AQUA / TERRA	NASA	Col. 5 Deep Blue Col. 6, 6.1	Aerosol total optical depth, fire radiative power	20090901 - 20150902 - 20170124 -
PMAp	METOP-A METOP-B	EUMETSAT		AOD	20170124 - 20170926 -

2.1.1 The CAMS o-suite

The o-suite consists of the IFS-CB05 chemistry combined with the CAMS bulk aerosol model. The chemistry is described in Flemming et al. (2015) and Flemming et al. (2017), aerosol is described in Morcrette et al. (2009). The forecast length is 120 h. The o-suite data is stored under **expver '0001'** of **class 'MC'**. On 21 June 2016 the model resolution has seen an upgrade from T255 to T511, and forecasts are produced twice per day.
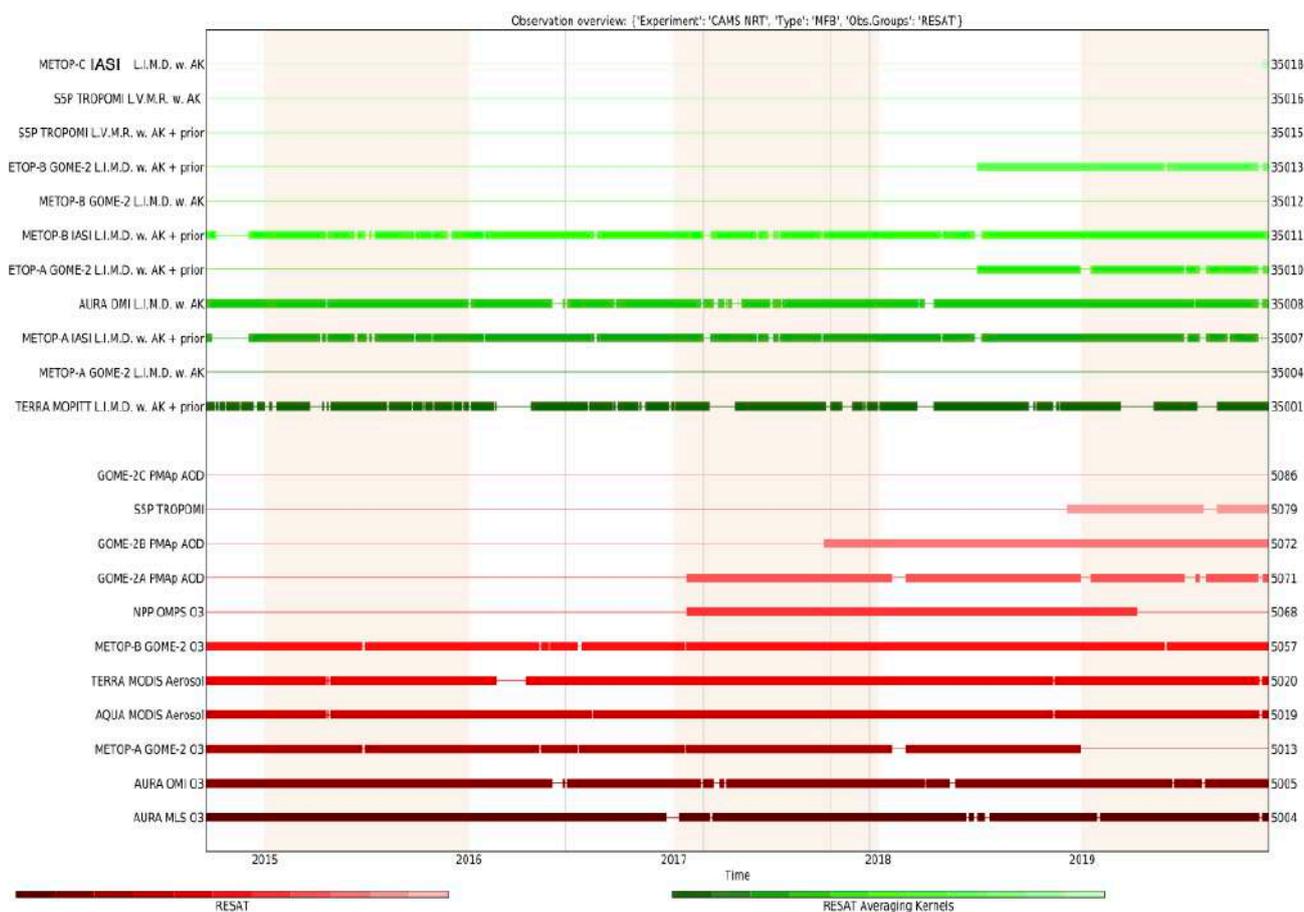
Figure 2.1: Satellite observation usage in the real-time analysis, for ozone, CO, aerosol AOD and NO_2, from October 2014 onwards. Top eight rows: products assimilated using averaging kernels. New assimilated products since the 24 January 2017 upgrade are the PMaP AOD including GOME-2B and OMPS ozone profile observations. Sentinel-5P TROPOMI ozone is assimilated since Dec. 2018 ($5079=O_3$) and other products are monitored ($35016=NO_2$, $35015=CO$, $5081=SO_2$).

A short summary of the main model specifications:

- The modified CB05 tropospheric chemistry is used (Williams et al., 2013), originally taken from the TM5 chemistry transport model (Huijnen et al., 2010)
- Stratospheric ozone during the forecast is computed from the Cariolle scheme (Cariolle and Teyssèdre, 2007) as already available in IFS, while stratospheric NOx is constrained through a climatological ratio of HNO_3/O_3 at 10 hPa.
- Monthly mean dry deposition velocities are based on the SUMO model provided by the MOCAGE team.
- Data assimilation is described in Inness et al. (2015) and Benedetti et al. (2009) for chemical trace gases and aerosol, respectively. Satellite data assimilated is listed in Table 2.2 and Fig. 2.1.
- Anthropogenic and biogenic emissions are based on MACCity (Granier et al., 2011) and a climatology of the MEGAN-MACC emission inventories (Sindelarova et al., 2014)
- NRT fire emissions are taken from GFASv1.2 (Kaiser et al. 2012).

The aerosol model includes 14 prognostic variables (Remy et al., 2019).

- 3 size bins each for sea-salt and desert dust

- 2 bins (hydrophibic and hydrophobic) each for organic matter and black carbon

- 1 bin for sulphate

- 2 bins (fine and coarse) for nitrate (New since 46R1)

- 1 bin for ammonium (New since 46R1)

The SO_2 precursor for sulphate aerosol no longer exists as a separate prognostic in the aerosol scheme, which since 46R1 couples directly to the SO_2 in the chemistry scheme instead. Likewise, the precursors for the new nitrate and ammonium aerosol (nitric acid and ammonia) are also part of the chemistry scheme rather than the aerosol scheme.

Aerosol total mass is constrained by the assimilation of MODIS and PMAp AOD (Benedetti et al. 2009). A variational bias correction is currently applied for the PMAp AOD based on the approach used also elsewhere in the IFS (Dee and Uppala, 2009).

A history of updates of the o-suite is given in Table 2.3, and is documented in earlier MACC-VAL and CAMS reports: https://atmosphere.copernicus.eu/node/326. This includes a list with changes concerning the assimilation system.

The CAMS o-suite system is upgraded regularly, following updates to the ECMWF meteorological model as well as CAMS-specific updates such as changes in chemical data assimilation. These changes are documented in e-suite validation reports, as can be found from the link above. Essential model upgrades are also documented in Table 2.3.

The penultimate upgrade of the system (45r1) took place on 26 June 2018. This upgrade is also relevant for this report (for the period up to 8 July), and the validation for this upgrade is described in Eskes et al., 2018b/2018c.

2.1.2 Short description of the latest CAMS upgrade (46r1)

The latest upgrade of the system took place on 9 July 2019 and is based on IFS version cy46r1_CAMS and involves the move from 60 to 137 vertical levels, see https://atmosphere.copernicus.eu/cycle-46r1 or https://confluence.ecmwf.int/display/COPSRV/Current+global+production+suites.

The validation for this 46r1 upgrade is described in Basart et al. 2019: https://atmosphere.copernicus.eu/sites/default/files/2019-07/CAMS84_2018SC1_D3.2.1-201907_esuite_v1.pdf

The meteorological changes can be found on the ECMWF-IFS CY46R1 page, https://confluence.ecmwf.int/display/FCST/Implementation+of+IFS+cycle+46R1.

The atmospheric composition content of the new cycle includes the following aspects:

Assimilation:

- New model-error covariance matrices for aerosol and chemistry at 137 levels.

Table 2.3: Long-term o-suite system updates.

Date	o-suite update
2009.08.01	Start of first NRT experiment f7kn with coupled MOZART chemistry, without aerosol. Also without data assimilation.
2009.09.01	Start of first MACC NRT experiment f93i, based on meteo cy36r1, MOZART v3.0 chemistry, MACC aerosol model, RETRO/REAS and GFEDv2 climatological emissions, T159L60 (IFS) and 1.875°×1.875° (MOZART) resolution.
2012.07.05	Update to experiment fnyp: based on meteo cy37r3, MOZART v3.5 chemistry, where changes mostly affect the stratosphere, MACCity (gas-phase), GFASv1 emissions (gas phase and aerosol), T255L60 (IFS) and 1.125°×1.125° (MOZART) resolution. Rebalancing aerosol model, affecting dust.
2013.10.07	Update of experiment fnyp from e-suite experiment fwu0: based on meteo cy38r2, no changes to chemistry, but significant rebalancing aerosol model. Assimilation of 21 layer SBUV/2 ozone product
2014.02.24	Update of experiment fnyp from e-suite experiment fzpr: based on meteo cy40r1. No significant changes to chemistry and aerosol models.
2014.09.18	Update to experiment g4e2: based on meteo cy40r2. In this model version IFS-CB05 is introduced to model atmospheric chemistry.
2015.09.03	Update to experiment g9rr: based on meteo cy41r1.
2016.06.21	Update to experiment 0067: based on meteo cy41r1, but a resolution increase from T255 to T511, and two production runs per day
2017.01.24	Update to cycle 43R1_CAMS, T511L60
2017.09.26	Update to cycle 43R3_CAMS, T511L60
2018.06.26	Update to cycle 45R1_CAMS, T511L60
2019-07-09	Update to cycle 46R1_CAMS, T511L137

Observations:

- No new atmospheric composition observations compared to Cycle 45r1.

Emissions:

- New emissions inventories: CAMS_GLOB_ANT v2.1 (anthropogenic) and CAMS_GLOB_BIO v1.1 (biogenic), in place of previous MACCity and MEGAN_MACC inventories.

- Biomass-burning injection heights from GFAS and updated diurnal cycle. In particular, this reduces the overestimation of near-surface PM2.5 during fire events.

- Anthropogenic SOA production was updated with a diurnal cycle and a regionally-varying ratio to CO emissions. This has a small impact on AOD, but significantly reduces night-time near-surface PM2.5 in polluted regions.

- New online dust emission scheme (Nabat et al., 2012). This increases total dust emissions and shifts them towards larger particle sizes, in line with recent literature. An updated dust source function improves the selection of source regions, reducing "gaps" in dust emissions.

- Sea-salt production over freshwater lakes eliminated. This corrects an issue that was particularly noticeable over the Great Lakes.

Other model changes:

- Vertical resolution increased from 60 levels to 137 levels, matching that used at ECMWF for NWP. This includes moving the model top from 0.1 hPa to 0.01 hPa.

- New nitrate and ammonium aerosol species are included and are coupled to the gas-phase nitrogen chemistry. This is a major expansion of the aerosol species represented in the model, giving a more complete representation of the species which contribute to e.g. PM2.5 over Europe.

- Sulphur species (SO_2 and SO_4) coupled between chemistry and aerosol schemes. See discontinued parameters below. This brings a greater consistency between the chemistry and aerosol products related to the sulphur cycle.

- Online calculation of dry deposition velocities for trace gases. This was already in place for aerosols in 45r1 and allows the deposition scheme to better account for variations in surface properties.

- Updates to wet deposition parameterisations. This brings improvements in the distinction between scavenging by liquid and ice and harmonises the treatment for aerosols and trace gases.

- Updates to chemical reaction rates following latest recommendations by JPL/IUPAC.

2.1.3 Control

The control run (relevant expver = **gzhy,** since 26/06/2018; expver = **h7c4** since 09/07/2019) applies the same settings as the respective o-suites, based on the coupled IFS-CB05 system with CAMS aerosol for cy54r1, except that data assimilation is not switched on. The meteorology in the control run is initialized with the meteorological fields from the o-suite.

2.1.4 High-resolution CO_2 and CH_4 forecasts and delayed-mode analyses

The pre-operational forecasts of CO_2 and CH_4 use an independent setup of the IFS at a resolution of TL1279, i.e. ~16 km horizontal, and with 137 levels. This system runs in real time and does not apply data assimilation for the greenhouse gases.

The land vegetation fluxes for CO_2 are modelled on-line by the CTESSEL carbon module (Boussetta et al., 2013). A biogenic flux adjustment scheme is used in order to reduce large-scale biases in the net ecosystem fluxes (Agusti-Panareda, 2015). The anthropogenic fluxes are based on the annual mean EDGARv4.2 inventory using the most recent year available (i.e. 2008) with estimated and climatological trends to extrapolate to the current year. The fire fluxes are from GFAS (Kaiser et al., 2012). Methane fluxes are prescribed in the IFS using inventory and climatological data sets,

consistent with those used as prior information in the CH_4 flux inversions from Bergamaschi et al. (2009). The anthropogenic fluxes are from the EDGAR 4.2 database (Janssens-Maenhout et al, 2012) valid for the year 2008. The biomass burning emissions are from GFAS v1.2 (Kaiser et al., 2012). The high-resolution forecast experiments also included a linear CO scheme (Massart et al., 2015).

The experiments analysed in this report are:

- "**ghqy**" from March 2016. The initial conditions used in ghqy on 1st of March 2016 are from the GHG analysis (experiment gg5m). Furthermore, the meteorological analysis used to initialize the ghqy forecast changed resolution and model grid in March 2016. Note that the CO_2, CH_4 and linear CO tracers are free-running.

- "**gqpe**" (43R1) from January 2017, and "**gznv**" (45R1) from 1 December 2018 to present. It runs with a TCO1279 Gaussian cubic octahedral grid (equivalent to approximately 9km horizontal resolution). Note that the CO_2, CH_4 and linear CO tracers are initialized with the GHG analysis (gqiq) for CO_2 and CH_4 and the CAMS operational analysis for CO.

- The greenhouse gas analysis experiment runs on a TCO399 grid (equivalent to around 25km) and 137 vertical levels and is available from January 2017 ("**gqiq**", 43R1) and 1 December 2018 ("**gwx3**", 45R1). This experiment runs in delayed mode (4 days behind real time) and makes use of observations from TANSO-GOSAT (methane and CO_2) and MetOp-IASI (methane).

2.2 Other systems

2.2.1 BASCOE

The NRT analyses and forecasts of ozone and related species for the stratosphere, as delivered by the Belgian Assimilation System for Chemical ObsErvations (BASCOE) of BIRA-IASB (Lefever et al., 2014; Errera et al., 2008), are used as an independent model evaluation of the CAMS products. The NRT BASCOE product is the ozone analysis of Aura/MLS-SCI level 2 standard products, run in the following configuration (version 05.07):

- The following species are assimilated: O_3, H_2O, HNO_3, HCl, $HOCl$, N_2O and ClO.
- It lags by typically 4 days, due to latency time of 4 days for arrival of non-ozone data from Aura/MLS-SCI (i.e. the scientific offline Aura/MLS dataset).
- Global horizontal grid with a 3.75° longitude by 2.5° latitude resolution.
- Vertical grid is hybrid-pressure and consists in 86 levels extending from 0.01 hPa to the surface.
- Winds, temperature and surface pressure are interpolated in the ECMWF operational 6-hourly analyses.
- Time steps of 20 minutes, output every 3 hours

See the stratospheric ozone service at http://www.copernicus-stratosphere.eu/. It delivers graphical products dedicated to stratospheric composition and allows easy comparison between the results of o-suite, BASCOE and TM3DAM. The BASCOE data products (HDF4 files) are also distributed from this webpage. Other details and bibliographic references on BASCOE can be found at http://bascoe.oma.be/. A detailed change log for BASCOE can be found at http://www.copernicus-stratosphere.eu/4_NRT_products/3_Models_changelogs/BASCOE.php.

2.2.2 TM3DAM and the multi-sensor reanalysis

One of the MACC products was a 30-year reanalysis, near-real time analysis and 10-day forecast of ozone column amounts performed with the KNMI TM3DAM data assimilation system, the Multi-Sensor Reanalysis (MSR) system (van der A et al., 2010, 2013),
http://www.temis.nl/macc/index.php?link=o3_msr_intro.html.
The corresponding validation report can be found at
http://www.copernicus-atmosphere.eu/services/gac/global_verification/validation_reports/.

The NRT TM3DAM product used for the validation of the CAMS NRT streams is the ozone analysis of Envisat/SCIAMACHY (until April 2012), AURA/OMI, and MetOp-A/GOME-2, run in the following configuration:

- total O_3 columns are assimilated
- Global horizontal grid with a 3° longitude by 2° latitude resolution.
- Vertical grid is hybrid-pressure and consists in 44 levels extending from 0.1 hPa to 100 hPa.
- Dynamical fields from ECMWF operational 6-hourly analysis.

An update of the MSR (MSR-2) was presented in van der A et al. (2015), which extended the record to 43 years based on ERA-interim reanalysis meteo and with an improved resolution of 1x1 degree.

2.2.3 SDS-WAS multimodel ensemble

The World Meteorological Organization's Sand and Dust Storm Warning Advisory and Assessment System (WMO SDS-WAS) for Northern Africa, Middle East and Europe (NAMEE) Regional Center (http://sds-was.aemet.es/) has established a protocol to routinely exchange products from dust forecast models as the basis for both near-real-time and delayed common model evaluation. Currently, twelve regional and global models (see the complete list in the following link https://sds-was.aemet.es/forecast-products/forecast-evaluation/model-inter-comparison-and-forecast-evaluation/at_download/file) provides daily operational dust forecasts (i.e. dust optical depth, DOD, and dust surface concentration).

Different multi-model products are generated from the different prediction models. Two products describing centrality (multi-model median and mean) and two products describing spread (standard deviation and range of variation) are daily computed. In order to generate them, the model outputs are bi-linearly interpolated to a common grid mesh of 0.5º x 0.5º. The multimodel DOD (at 550 nm) Median from nine dust prediction models participating in the SDS-WAS Regional Center is used for the validation of the CAMS NRT streams.

2.3 CAMS products

An extended list of output products from the NRT stream o-suite are available as 3-hourly instantaneous values up to five forecast days. These are available from ECMWF (through ftp in grib2 and netcdf format, https://atmosphere.copernicus.eu/data).

2.4 Availability and timing of CAMS products

The availability statistics provided in Table 2.6 are computed for the end of the 5-day forecast run. The CAMS production KPI is defined as the percentage of cycles in which all the general data

dissemination tasks are completed before the deadlines: 10 UTC for the 00:00 and 22 UTC for the 12:00 UTC run. This was in part based on requirements from the regional models. We note that at present most regional models can still provide their forecasts even if the global forecast is available a bit later. Note that since 21 June 2016 two CAMS forecasts are produced each day.

The o-suite data delivery for the period September-November 2019 (SON-2019) was good, with an on-time percentage of 98.90 %. There were two small delays in September due to problems with HPC.

See table 2.6 for detailed statistics from 2014 to SON-2019.

Table 2.6: Timeliness of the o-suite from December 2014. From June 2016 onwards CAMS has produced two forecasts per day.

Months	On time, 10 & 22 utc	80th perc	90th perc	95th perc
Dec-Feb '14-'15	97%	D+0, 19:43	D+0, 20:28	D+0, 21:13
Mar-May 2015	96%	D+0, 19:38	D+0, 21:03	D+0, 21:40
Jun-Aug 2015	95%	D+0, 20:24	D+0, 20:53	D+0, 21:54
Sept-Nov 2015	95%	D+0, 19:44	D+0, 20:55	D+0, 21:51
Dec-Feb '15-'16	100%	D+0, 18:39	D+0, 18:57	D+0, 19:43
Mar-May 2016	98%	D+0, 19:32	D+0, 19:47	D+0, 20:00
Jun-Aug 2016 (00 and 12 cycle)	100%	D+0, 08:53 D+0, 20:55	D+0, 09:04 D+0, 21:01	D+0, 09:18 D+0, 21:18
Sep-Nov 2016	98.9%	D+0, 08:44 D+0, 20:44	D+0, 08:51 D+0, 20:48	D+0, 08:52 D+0, 20:51
Dec 2016 - Feb 2017	99.4%	D+0, 09:02 D+0, 21:01	D+0, 09:11 D+0, 21:02	D+0, 09:18 D+0, 21:04
Mar-May 2017	100%	D+0, 09:08 D+0, 21:07	D+0, 09:14 D+0, 21:09	D+0, 09:19 D+0, 21:11
Jun-Aug 2017	100%	D+0, 09:05 D+0, 21:05	D+0, 09:07 D+0, 21:08	D+0, 9:09 D+0, 21:10
Sep-Nov 2017	100%	D+0, 09:02 D+0, 21:00	D+0, 09:05 D+0, 21:04	D+0, 9:09 D+0, 21:07
Dec 2017 - Feb 2018	98.33%	D+0, 08:55 D+0, 20:54	D+0, 08:59 D+0, 20:59	D+0, 09:01 D+0, 21:02
Mar-May 2018	98.9%	D+0, 09:00 D+0, 21:00	D+0, 09:06 D+0, 21:03	D+0, 09:08 D+0, 21:06
Jun-Aug 2018	100%	D+0, 09:11 D+0, 21:07	D+0, 09:14 D+0, 21:09	D+0, 09:20 D+0, 21:11
Sep-Nov 2018	100%	D+0, 09:05 D+0, 21:03	D+0, 09:09 D+0, 21:07	D+0, 09:13 D+0, 21:10
Dec 2018 - Feb 2019	98.85%	D+0, 09:03 D+0, 21:04	D+0, 09:06 D+0, 21:06	D+0, 09:08 D+0, 21:10
Mar-May 2019	100%	D+0, 09:07 D+0, 21:05	D+0, 09:10 D+0, 21:09	D+0, 09:12 D+0, 21:11
Jun-Aug 2019	99.46%	D+0, 09:19 D+0, 21:14	D+0, 09:22 D+0, 21:17	D+0, 09:27 D+0, 21:19
Sep-Nov 2019	98.90%	D+0, 09:14 D+0, 21:07	D+0, 09:23 D+0, 21:20	D+0, 09:26 D+0, 21:24

3. Tropospheric Ozone

3.1 Validation with sonde data in the free troposphere

Model profiles of the CAMS runs were compared to free tropospheric balloon sonde measurement data of 38 stations taken from the NDACC, WOUDC, NILU and SHADOZ databases for January 2013 to November 2019 (see Fig. 3.1.1 - 3.1.2). Towards the end of the period, the number of available soundings decreases, which implies that the evaluation results may become less representative. The figures contain the number of profiles in each month that are available for the evaluation. The methodology for model comparison against the observations is described in Douros et al., 2017. The free troposphere is defined as the altitude range between 750 and 200hPa in the tropics and between 750 and 300hPa elsewhere.

Please note that recent scientific findings (https://tropo.gsfc.nasa.gov/shadoz/Archive.html, Thompson et al., 2017; Witte et al., 2017; 2018, Stauffer, et al. in preparation 2020) show a drop-off in Total Ozone at various global ozone stations in comparison with satellite instruments. This drop-off amounts between 5-10% for stratospheric ozone. Changes in the ECC ozone instrument are associated with the drop-off, but no single factor has been identified as cause yet. For tropospheric ozone (<50 hPa) no alternations are reported, but cannot be ruled out. Data availability is thus recently limited.

MNMBs for the o-suite are mostly within the range ±20%, for all months, in all zonal bands, except for the Tropics and Antarctica, where larger positive MNMBs up to ±45% appear, see Fig. 3.1.4. During the last year (November 2018 to November 2019) MNMBs are ±20% over the Arctic and Northern Midlatitudes and up to 30% for Antarctica and the Tropics, see Fig. 3.1.1.-3.1.4.

Over the Arctic, the o-suite mostly shows slightly negative MNMBs during winter and spring (MNMBs up to -13%) and positive biases during summer and fall (MNMBs up to 19%) see, Fig. 3.1.1.

Over the NH mid-latitudes MNMBs for the o-suite are on average close to zero and only during the summer season maxima up to 12% appear. The control run shows comparable results, except for negative biases up to -10% during the summer season.
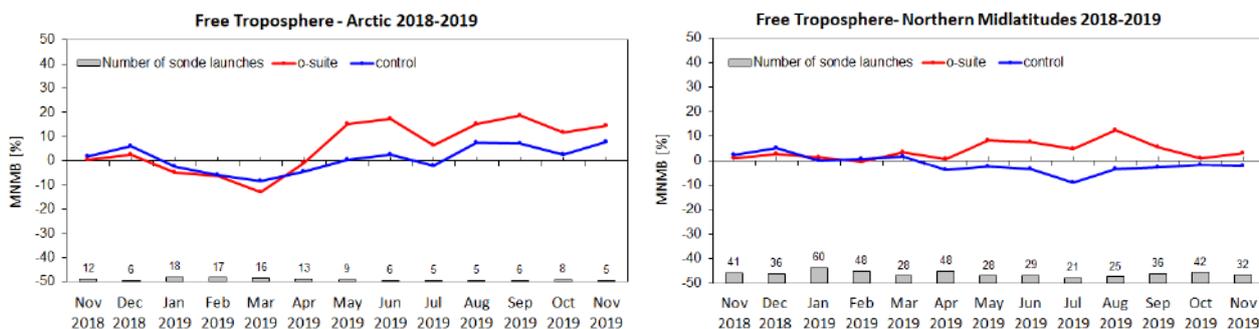

Figure 3.1.1: MNMBs (%) of ozone in the free troposphere (between 750 and 300 hPa) from the IFS model runs against aggregated sonde data over the Arctic (left) and the Northern mid latitudes (right). The numbers indicate the amount of individual number of sondes.
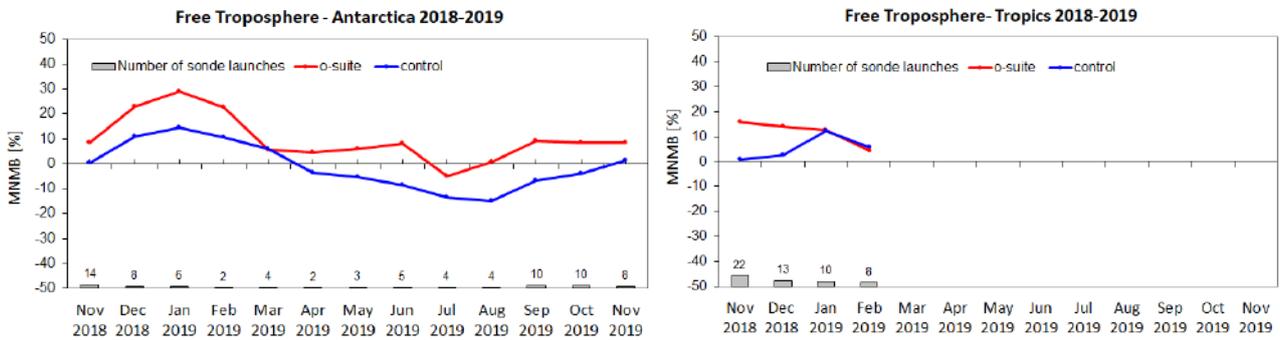
Figure 3.1.2: MNMBs (%) of ozone in the free troposphere (between 750 and 200hPa (Tropics) / 300hPa) from the IFS model runs against aggregated sonde data over the Tropics (left) and Antarctica (right). The numbers indicate the amount of individual number of sondes.
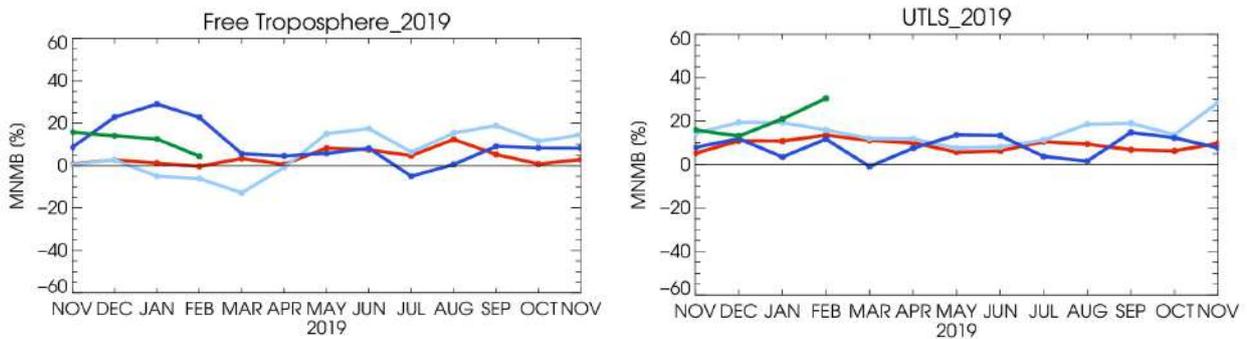

Figure 3.1.3: MNMBs (%) of ozone in the free troposphere (left, between 750 and 200hPa (Tropics) / 300hPa) and UTLS (right, between 300 and 100hPa (Tropics) / 60hPa) from the IFS model runs against aggregated sonde data over the Tropics (green) and Antarctica (blue), Arctic (light blue) and Northern Midlatitudes (red).
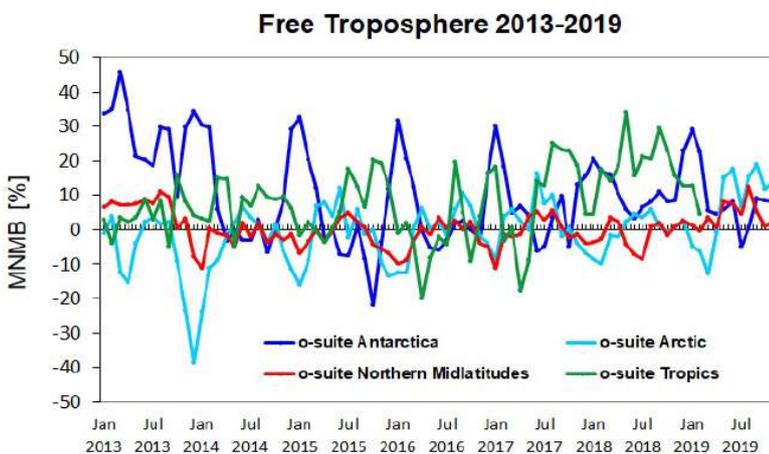
Figure 3.1.4: Time series of MNMB of ozone in the o-suite, compared against ozone sondes, averaged over different latitude bands. The free troposphere is defined here as the layer between 750 and 300 hPa.

Over Antarctica, ozone mixing ratios are mostly overestimated by the o-suite (up to 30%) by the o-suite, see Fig. 3.1.2. The control run shows negative MNMBs up to -15%.

In the UTLS, ozone is overestimated by the o-suite over all regions. MNMBs range within ±20%.
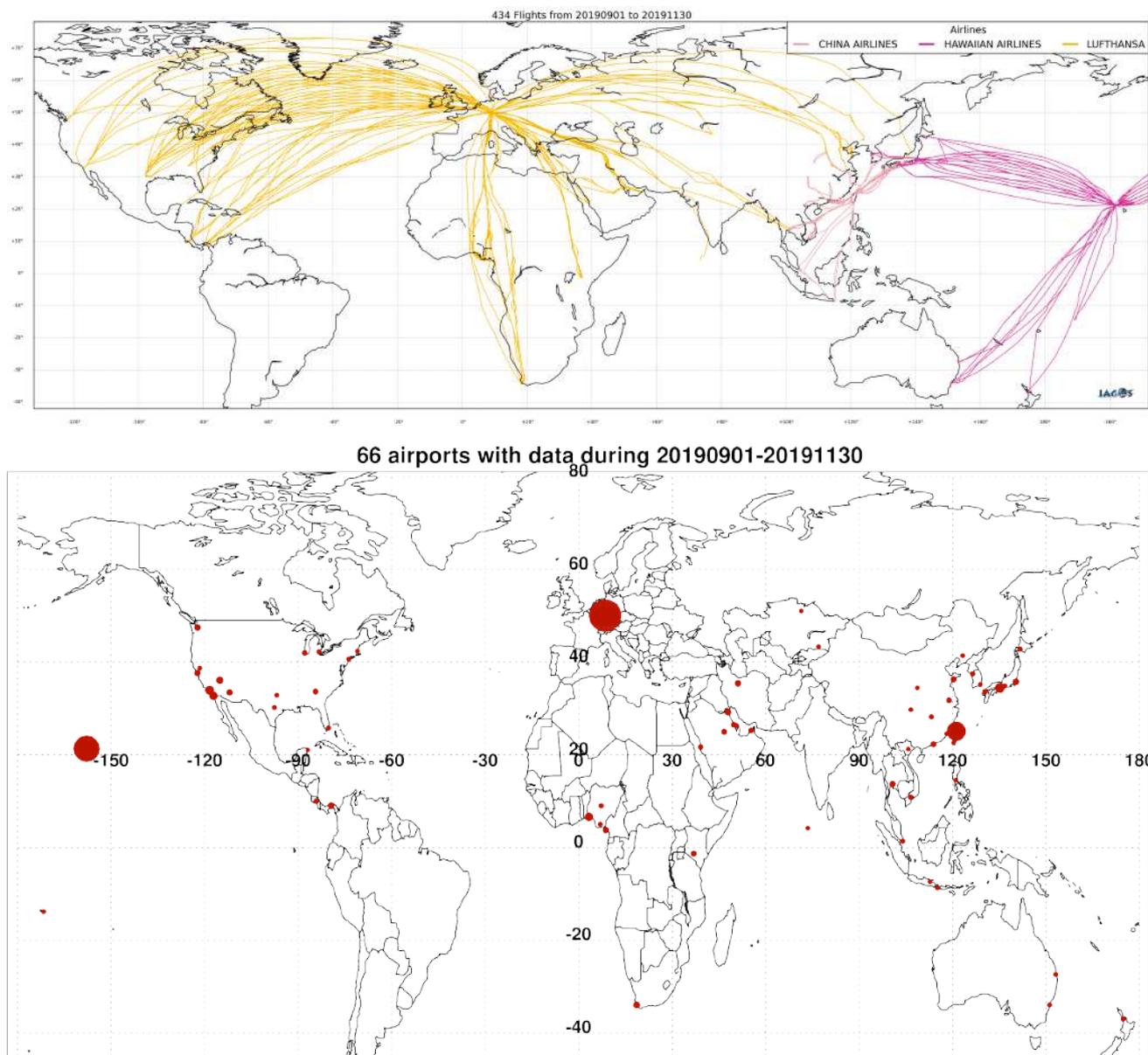
Figure 3.2.1. Map of the flights (top) and the visited airports (bottom) during the period September - November 2019, by the IAGOS-equipped aircraft. The size of the plotting circle represents the number of profiles available.

3.2 Ozone validation with IAGOS data

The daily profiles of ozone measured at airports around the world are shown on the CAMS website at http://www.iagos-data.fr/cams/nrt_profiles.php. For the period from September - November 2019, the data displayed on the web pages and in this report include only the data as validated by the instrument PI. The available flights and available airports are shown in Fig. 3.2.1 top and bottom respectively. Performance indicators have been calculated for different parts of the IAGOS operations.
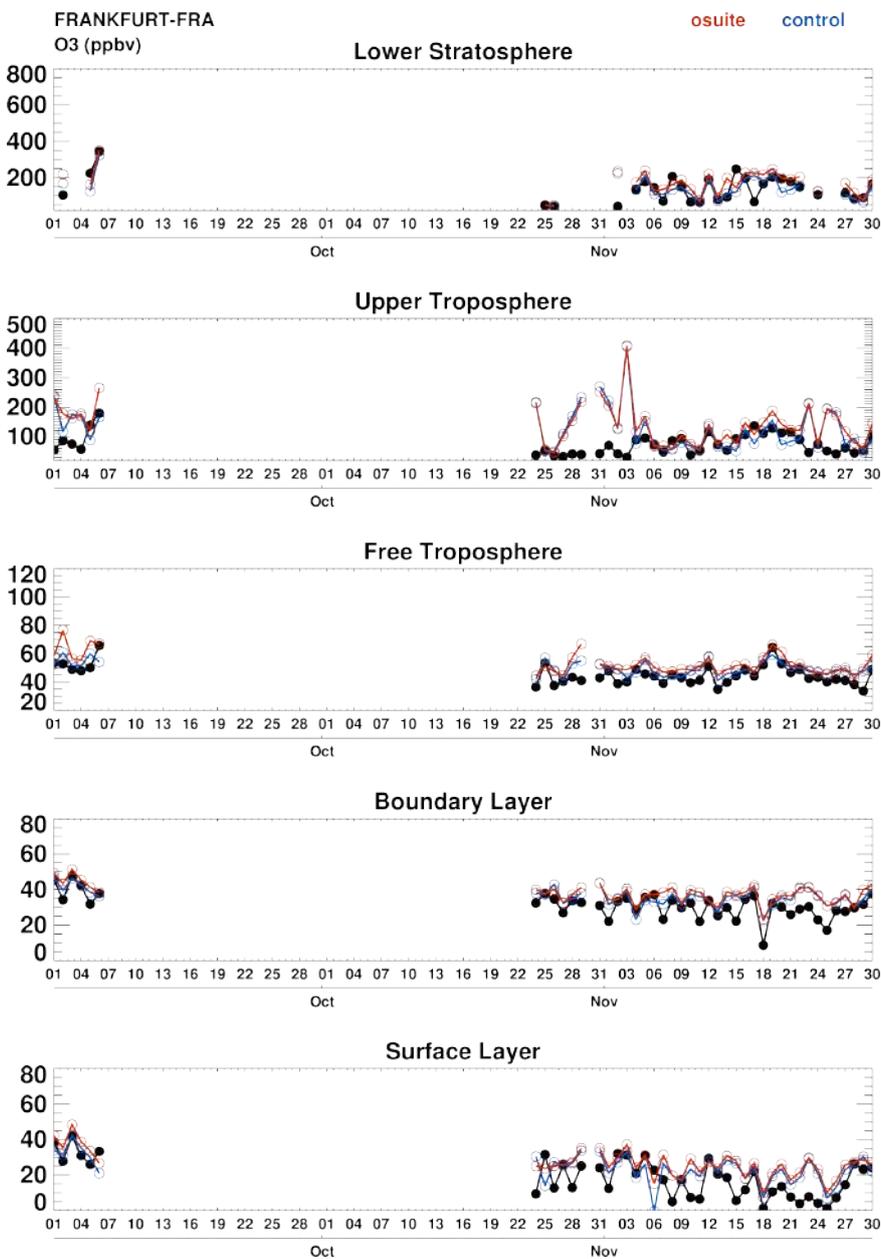
Figure 3.2.2. Time series of daily mean ozone over Frankfurt during SON 2019 for 5 layers: Surface Layer, Boundary Layer, Free Troposphere, Upper Troposphere and Lower Stratosphere. The o-suite is shown in red and associated control run in blue.

Six aircraft were operating during this period. With these aircrafts, operating fully over the three-month period, we can expect a total of about 1260 flights. The actual number of flights within the period was 434 (868 profiles) giving a performance of 34 %. These flights are shown in Fig. 3.2.1 (top). Twenty three percent (23%) of the operational flights had usable measurements of ozone and 8% of the flights had usable CO. Delivering these O_3 and CO data were two aircraft from Lufthansa operating from Frankfurt, and one from China Airlines. Fig. 3.2.1 (bottom) shows the available airports, with a plotting circle scaled to the highest number of flights at an airport.
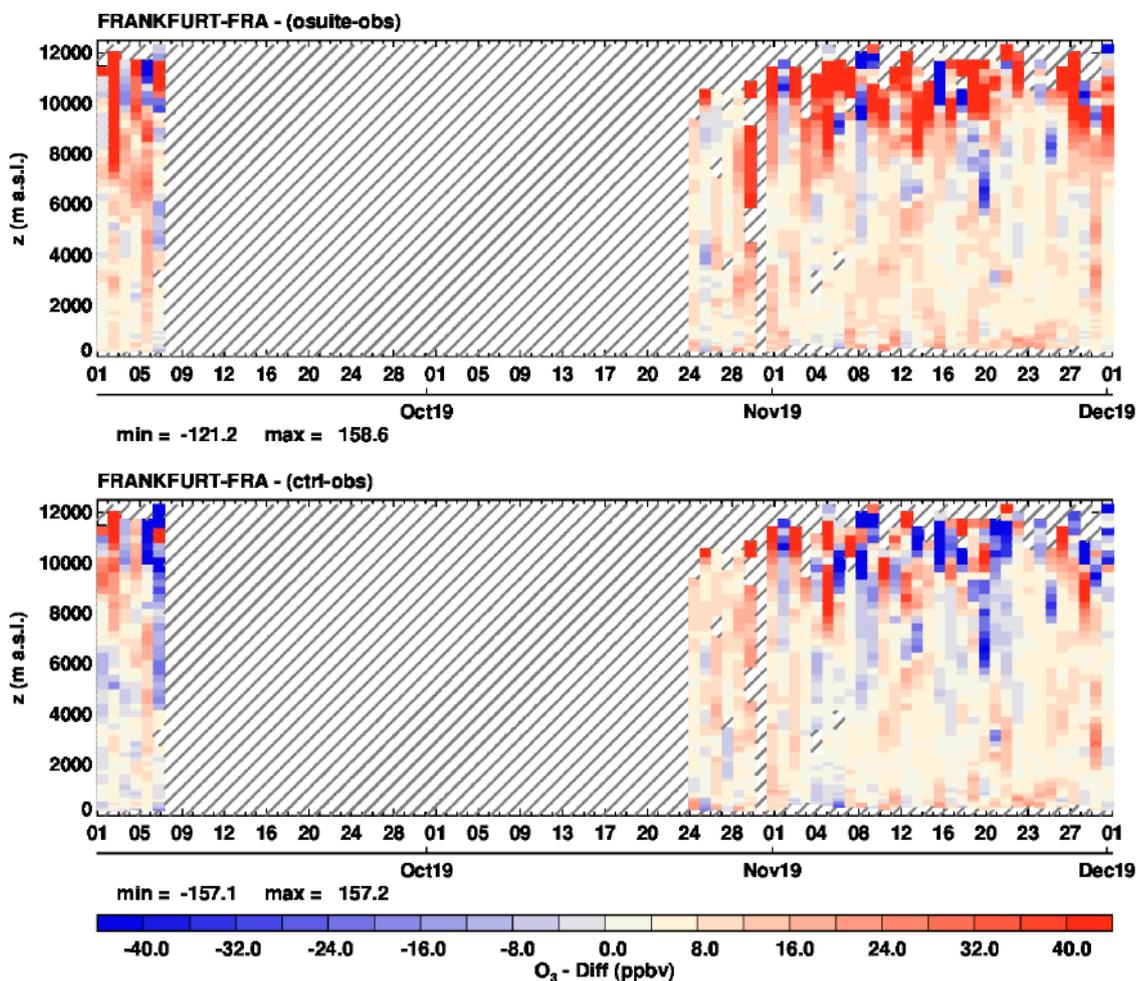
Figure 3.2.3 Time series of the absolute differences (model - observations) in daily profiles for ozone over Frankfurt during SON 2019. The top panel corresponds to o-suite the bottom panel to control run.

Europe

Fig. 3.2.2 presents ozone time series at Frankfurt during the full period September – November 2019 for 5 atmospheric layers. Time series of the profile differences (in ppbv) are also presented in Fig. 3.2.3. As it can be seen, no data is available between 7 September and 24 October, which is due a problem with the transmission from one of the packages. Observations for this missing period might be available afterward but with some more delay. As compared to the two previous reports, the bias in ozone is clearly smaller, which further confirms the reliability of IAGOS observations since the replacement of one of the Lufthansa instruments in mid-August 2019 (see MAM and JJA 2019 reports). In the low to mid-troposphere the two runs behave similarly. Ozone is well represented by both runs in the low troposphere with mostly small overestimations, and the best agreement is obtained in the free troposphere. In the UTLS region the bias is larger and the results from the two runs differ, ozone is mostly overestimated by the o-suite whereas the control run often underestimates ozone. Some examples of individual profiles at Frankfurt are also presented in Fig. 3.2.4.a-b.
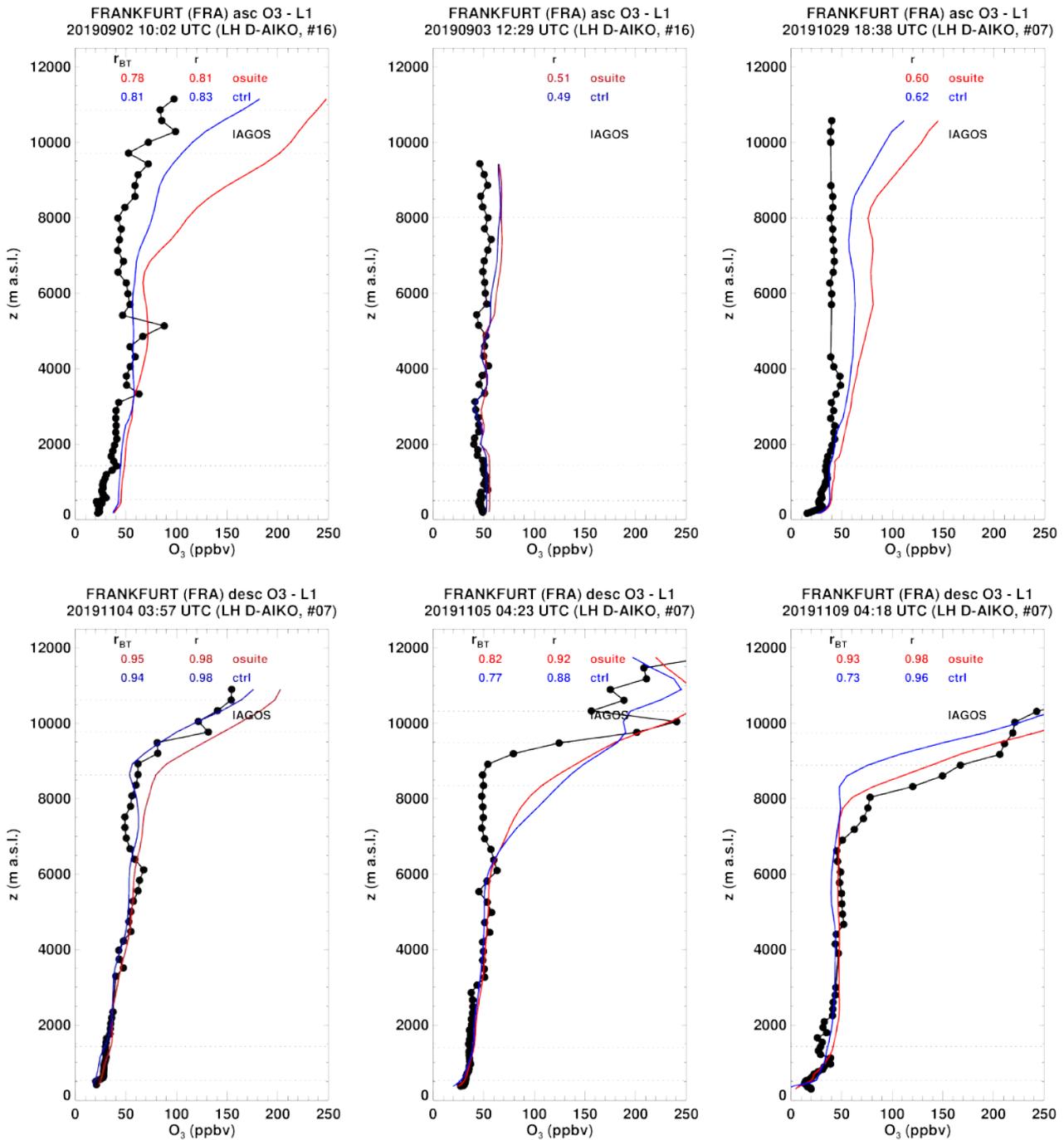
Figure 3.2.4.a Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Europe during SON 2019.

As shown in the time series, no major episode of ozone is observed during this period as ozone values remain below 50 ppbv in the surface and boundary layer. In the presented profiles, these ozone values are in general well represented by the two runs and are sometimes overestimated by the models (Fig. 3.2.4). On the profiles of 20 November, a peak is observed in the free troposphere near 3000 m with a value of about 100 ppbv. This ozone maximum is detected by both runs but underestimated by the models, which depict a much smoother shape of the profile.
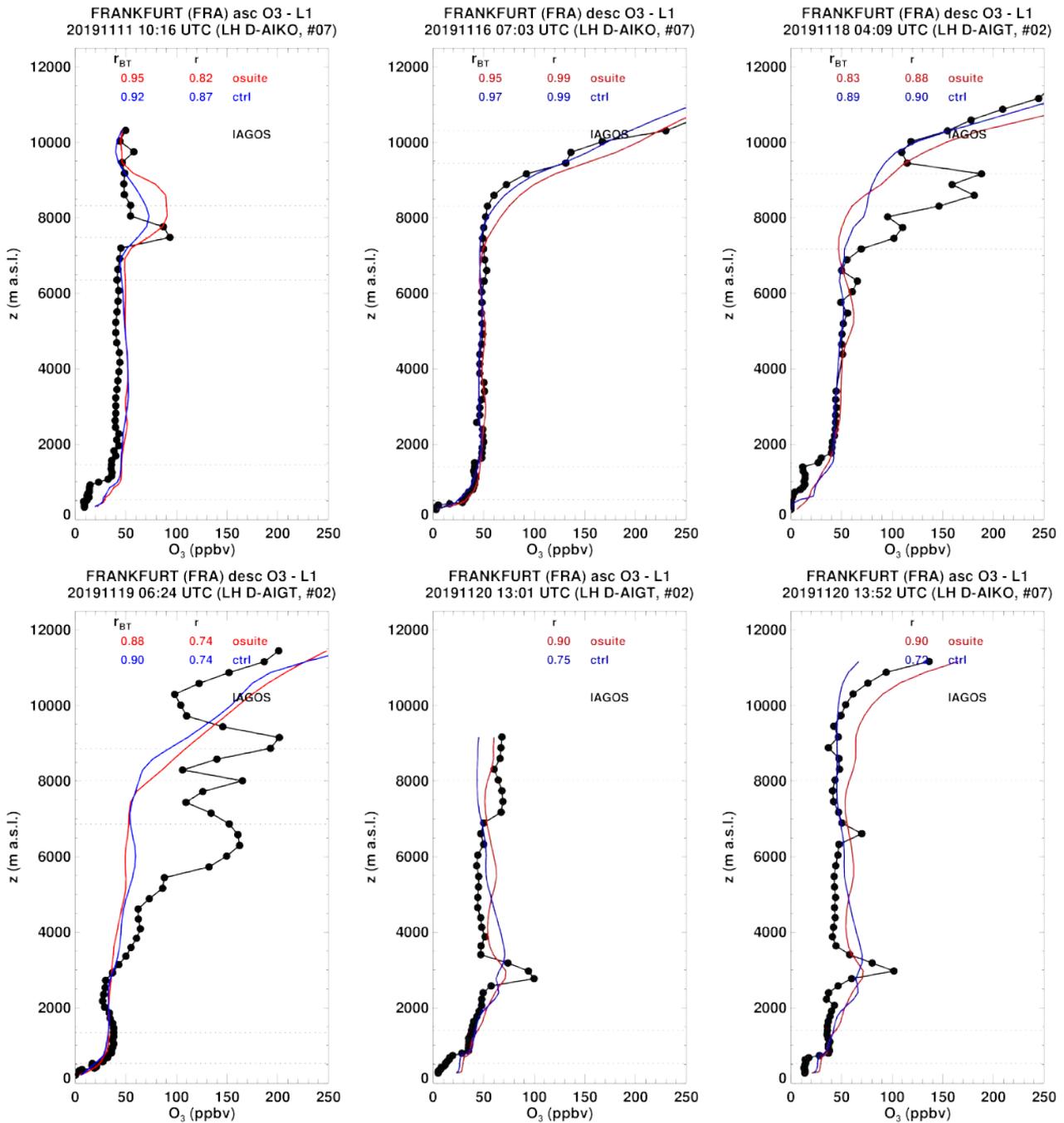
Figure 3.2.4.b Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Europe during SON 2019.

Another sharp peak in ozone of 90 ppbv is observed much higher at about 7500 m on 11 November. The latter is also detected by the models but slightly higher above, with a better performance in terms of magnitude from the o-suite as for the control run. On several of the profiles presented in Fig. 3.2.4, it can be seen that the model generally detects well tropopause descents (days 1104, 1105, 1109, 1116, 1118, 1119,), however the presence of ozone maxima near the tropopause are not reproduced by the models (days 1118 and 1119). When model and observations differ in the tropopause this leads to large biases like those observed in the timeseries in late October-early November (i.e. day 1029).

Middle East

Profiles are available over several airports across the Middle East, at the airports of Jeddah, Kuwait City, Riyadh, Bahrain, Dammam and Dubai (Fig. 3.2.5.a-b). For these profiles, ozone is in general in the range of 50-70 ppbv at the surface and boundary layer with the exception of a profile at Kuwait City on 2 November where the boundary layer presents an ozone peak of 110 ppb. In general, the results from both run are very similar with a good agreement in the surface and boundary layer. In the free troposphere and UTLS, both runs often overestimate ozone, however, the o-suite shows better results in these examples.

Africa

IAGOS profiles are available at airports in the gulf of Guinea at Port Harcourt, Lagos and Malabo, as well as at the Kenyan airport of Nairobi (Fig. 3.2.6.a-b). In most of the profiles, surface and boundary layer values are below 50 ppbv and in general both runs agree well with observations. Some of the profiles present maxima in the free troposphere (Lagos 0902, 0905, Malabo 0904, 1108, Nairobi 1103, 1116 and 1105), which are not always detected by the model. For most of the aforementioned profiles (except at Nairobi), CO profiles are also available (see CO section) and correlated to those of ozone and can therefore be related to the transport of pollution from forest fires occurring in Southern Africa during this period. These maxima are not always detected by the models, which behave similarly. For the two profiles at Lagos with a maximum at about 3500 meters, no maximum is found by the models. In the profiles at Malabo, the two runs reproduce the shape of the profiles with the correct magnitude of the maxima although the altitude of the maxima is often slightly different from those of the observations. Similar behaviour is found at Nairobi with maxima that are not always detected by the models.
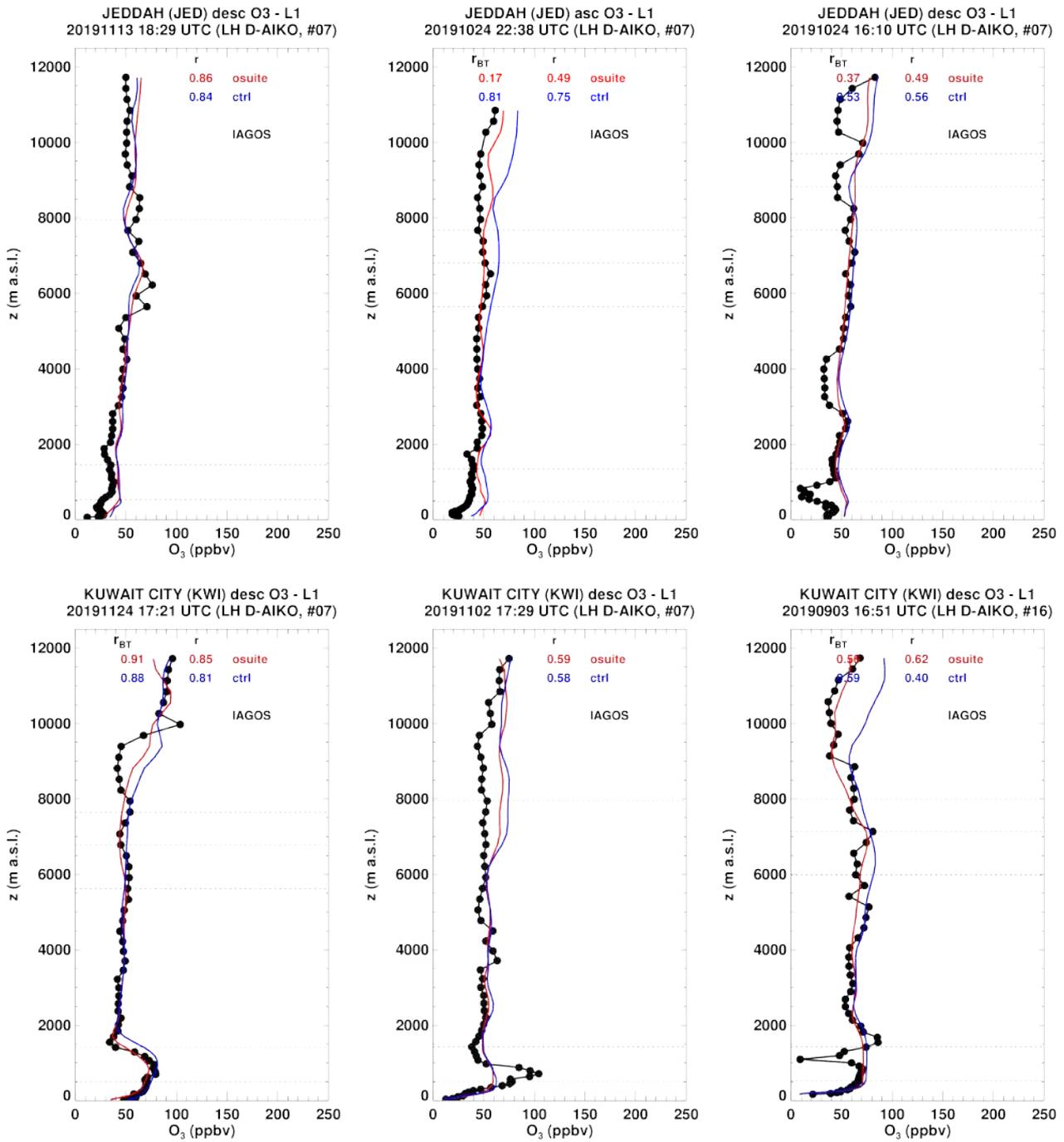
Figure 3.2.5.a Daily profile for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Middle East during SON 2019.
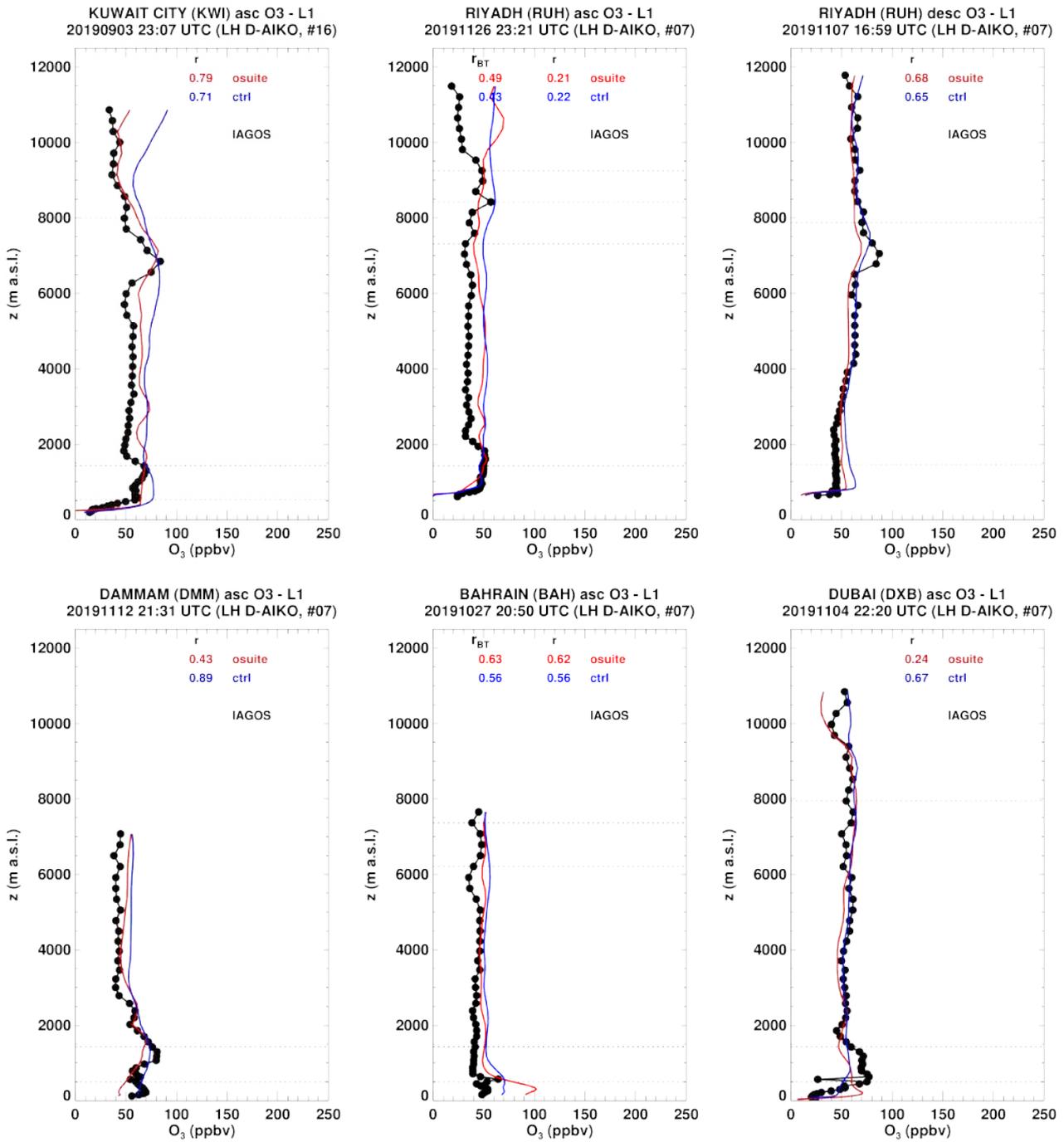
Figure 3.2.5.b Daily profile for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Middle East during SON 2019.
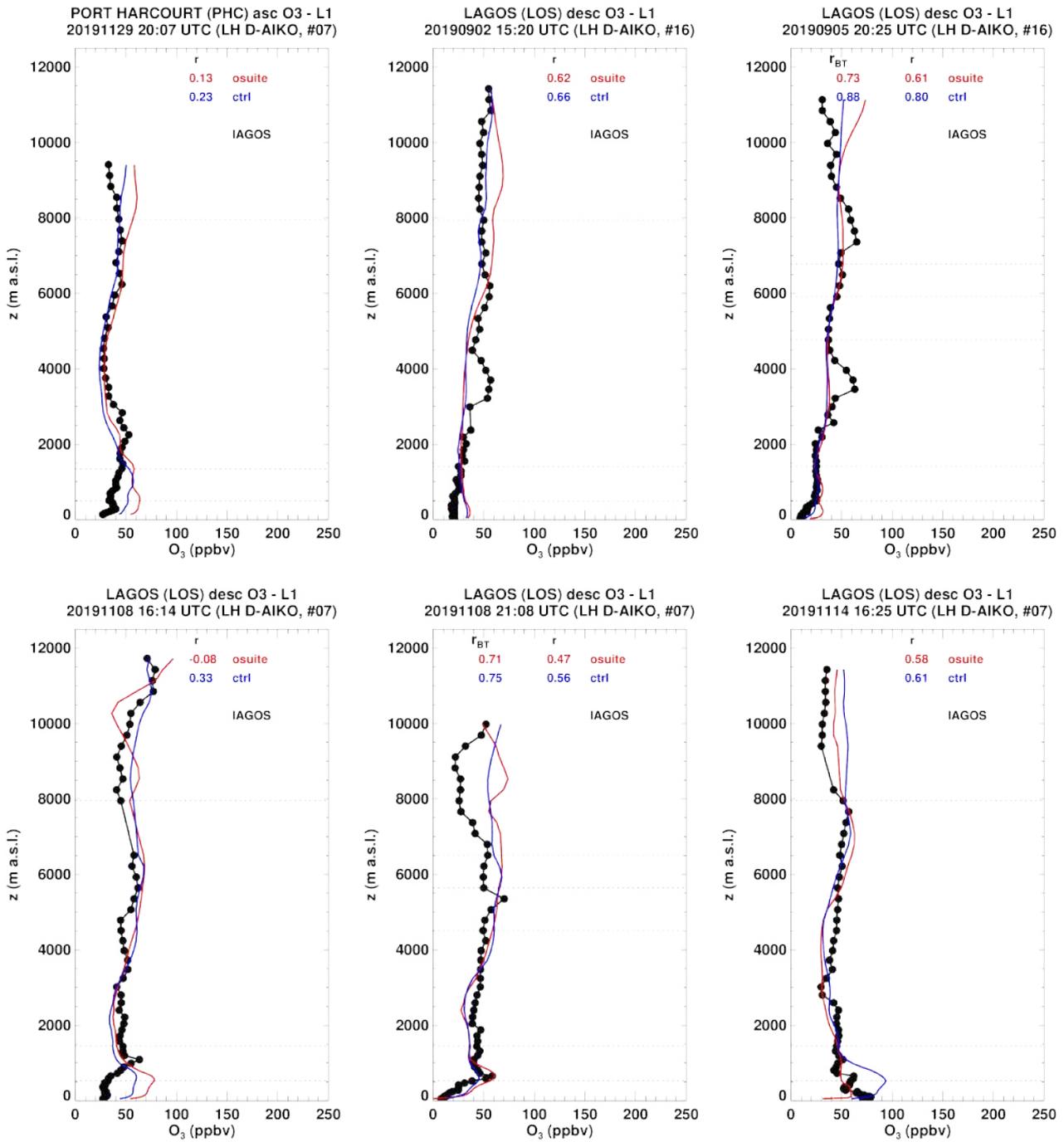
Figure 3.2.6.a Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over West Africa during SON 2019.
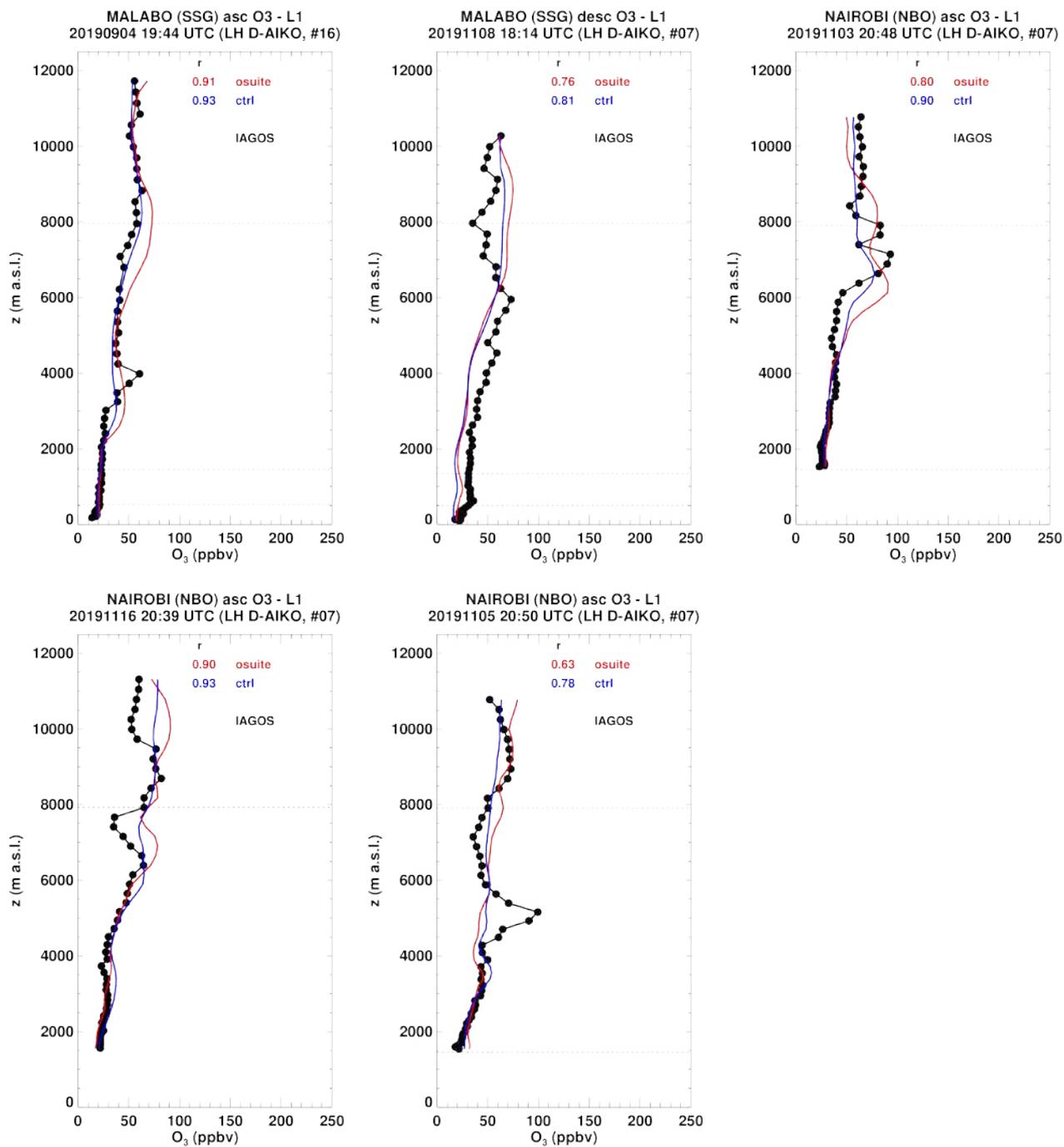
Figure 3.2.6.b Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over West Africa during SON 2019.
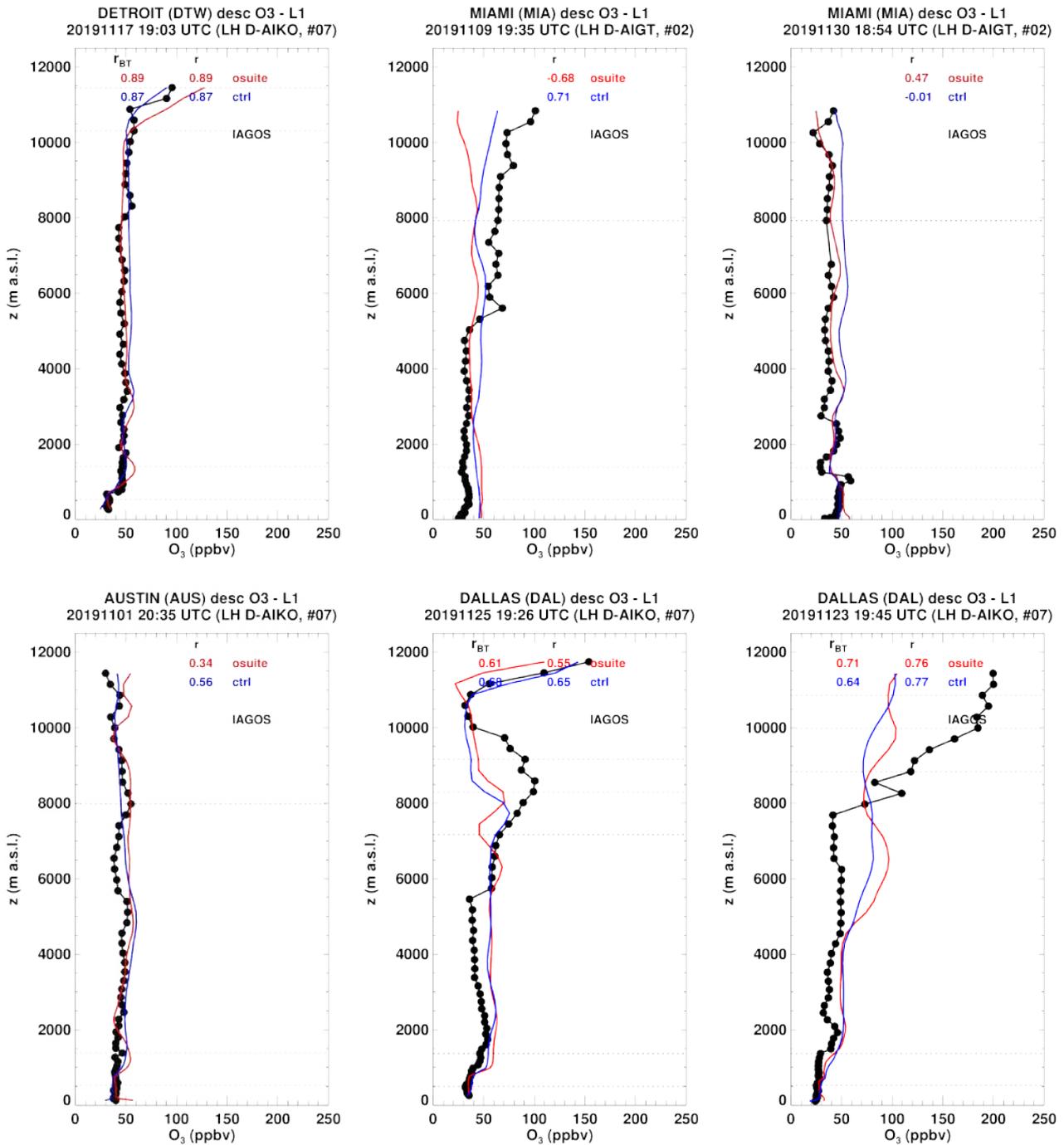
Figure 3.2.7.a Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during SON 2019.
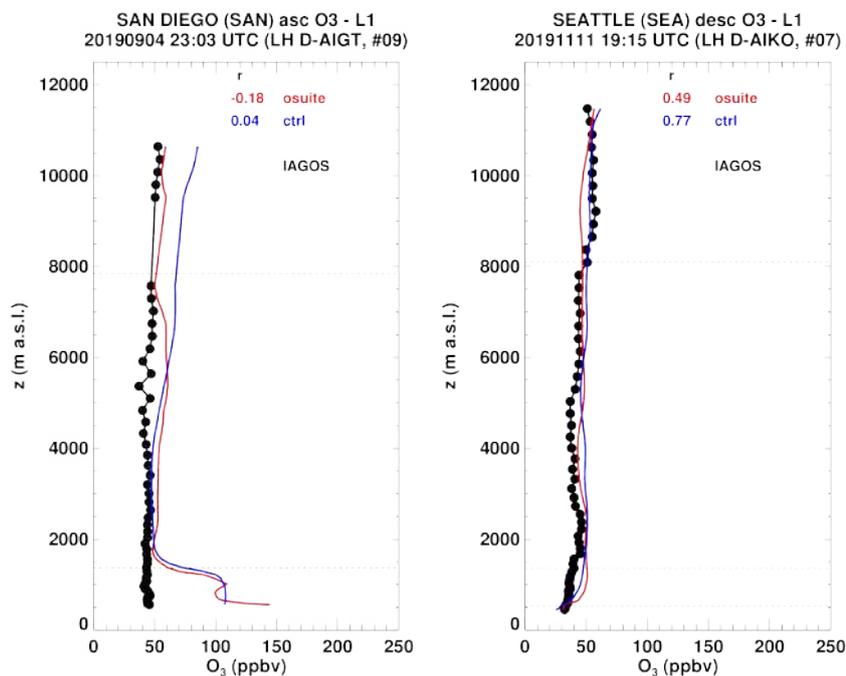
Figure 3.2.7.b Daily profiles for ozone from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during SON 2019.

North America

In North America, IAGOS observations during SON 2019 are available for: Detroit, Miami, Dallas, Austin, San Diego and Seattle. For all profiles, ozone values are generally near or below 50 ppbv from the surface to the mid-troposphere (Fig. 3.2.7.a-b). In these layers the results from both runs are very similar with a rather good agreement except in San Diego where large overestimations are found at the surface and boundary layer. In the UTLS, the bias is larger than in the low troposphere for both runs. At Dallas on 25 November, a maximum of nearly 100 ppbv is observed at the altitude of about 9000 meters. Both runs detect this maximum, which is underestimated but at a slightly lower altitude especially for the control run. At the same airport on 23 November, both runs fail to depict the tropopause and the shape of the profile from the models is very different from that of observations.

East Asia

Over East Asia only two ozone profiles are available: one at Qingdao and the other at Shenyang (Fig. 3.2.8). These two observed profiles present nearly constant shapes from the surface to the free troposphere with ozone values around 50 ppbv. In these layers, results from the o-suite and control run are very similar except for the boundary layer for Qingdao, where the o-suite performs slightly better than the control run, which overestimates O_3. The runs start to differ in the UTLS and the agreement with observations is worse than in the low to mid-troposphere.
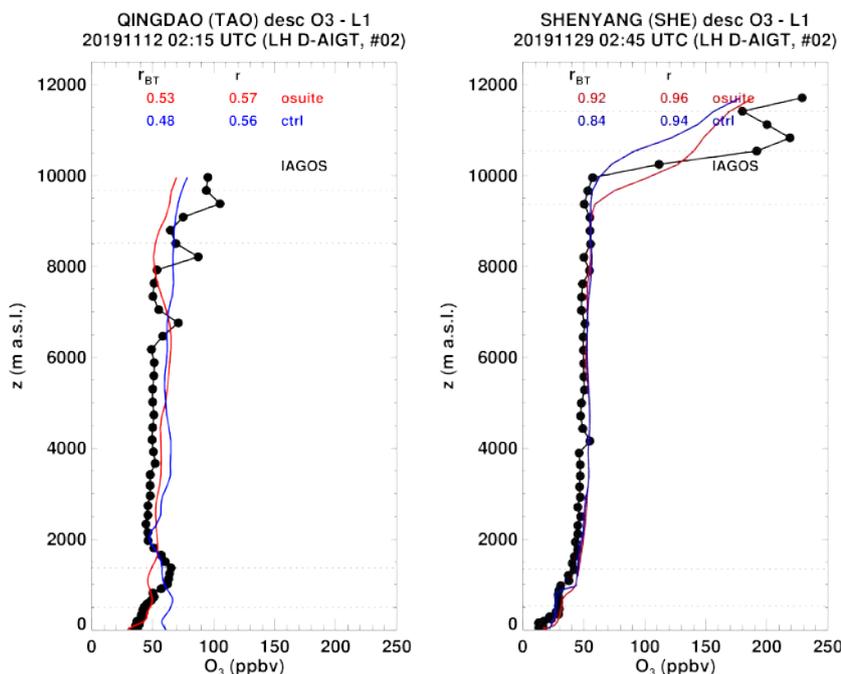
Figure 3.2.8. Daily profiles for ozone from IAGOS (black) and the two NRT products (o-suite: red, control: blue) over East Asia during SON 2019.
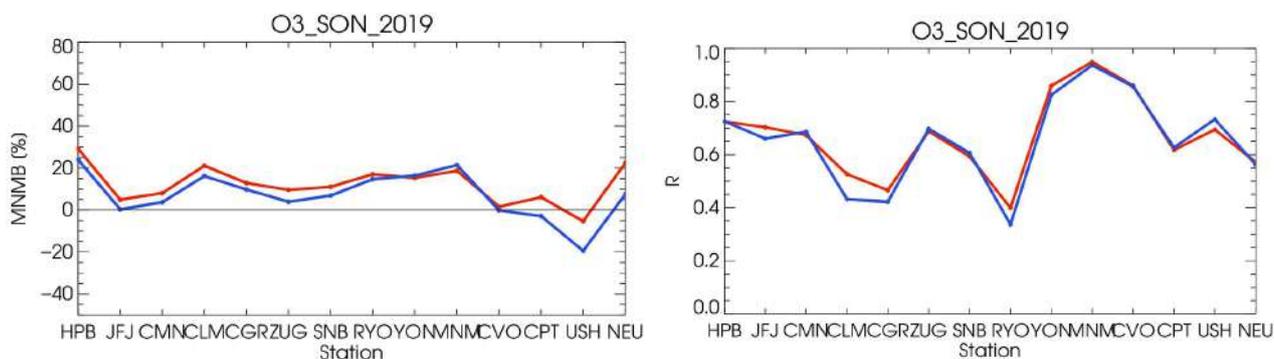

Figure 3.3.1: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT model runs compared to observational GAW data in the period September 2019 to November 2019 (o-suite: solid red; control: blue).

3.3 Validation with GAW and ESRL-GMD surface observations

For the Near Real Time (NRT) validation, 13 GAW stations and 14 ESRL stations are currently delivering O_3 surface concentrations in NRT, and the data are compared to model results. In the following, a seasonal evaluation of model performance for the 2 NRT runs (o-suite and control) has been carried out for the period from September to October 2019. The latest validation results based on GAW stations and based on ESRL observations can be found on the CAMS website, see section 1, table 1.2.
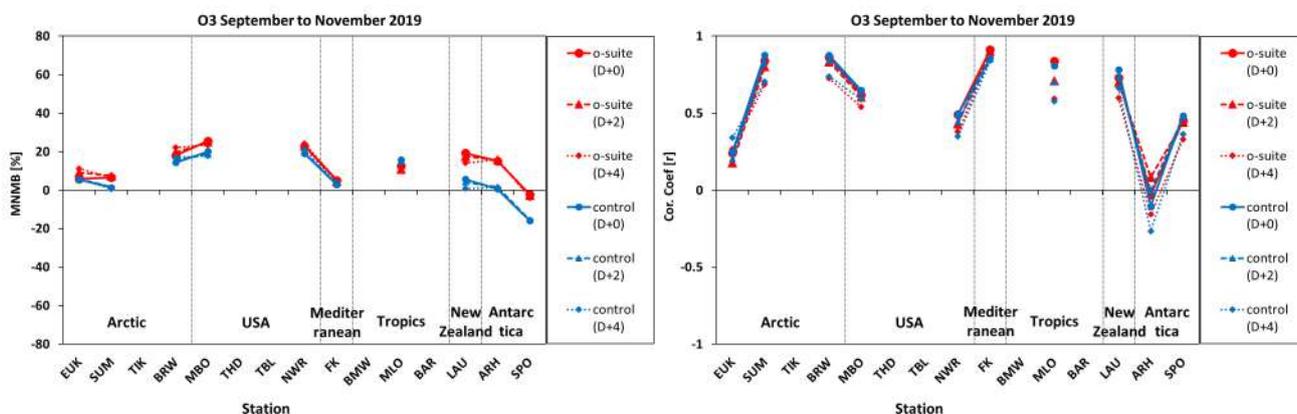
Figure 3.3.2: Modified normalized mean bias in % (left) and correlation coefficient (right) of the NRT forecast runs compared to observational ESRL data in the period September-November 2019. Circles correspond to D+0, triangles to D+2 and rhombs to D+4 metrics respectively.

Modified normalized mean biases in % (left panel) and correlation coefficients (right panel) for different forecasts days (D+2, red-dashed and D+4, red- pointed) with respect to GAW and ESRL observations are shown in Figs. 3.3.1 and 3.3.2. It indicates that MNMBs for both o-suite and control run mostly remain stable up to D+4 (forecast run from 96h to 120h). Correlations between simulated and observed surface ozone values remain almost stable up to D+2 (forecast run from 48h to 72h), but then drop (correlations for D+4 are lower than correlations for D+2 and D+0), see Fig. 3.3.1 and 3.3.2, right graph).

A comparison of the seasonal-mean MNMB over Europe (Fig. 3.3.3) from December 2012 to present shows minimal MNMBs during the winter season and larger biases in other months. Also, on average the MNMB for the o-suite and control shows an improvement over the years. The temporal correlation is consistently better for the control run than for the o-suite, but the o-suite shows strong improvements recently. The GAW results are summarized in Figs 3.3.1 and 3.3.3.

Looking at different regions, for European stations (HPB, JFJ, ZUG, SNB, CMN, CLM, CGR), observed O_3 surface mixing ratios are very close to the observations (Fig. 3.3.4). MNMBs are between -5% and 30% for the o-suite and between -20% and 30% for the control run, see Fig. 3.3.1. Correlations for European stations are between 0.46 and 0.95 for the o-suite and between 0.42 and 0.94 for the control run, see Fig. 3.3.1.

Over Arctic stations (EUK, BRW and SUM), the o-suite overestimates surface ozone values between 5% (at Summit and at Eureka) and 15% (at Point Barrow) (Fig. 3.3.5). On the other hand, the control run's O_3 surface mixing ratios are closer to the observations (MNMBs between 1% at SUM and -11% at BRW). Correlations between modelled and observed ozone values for both runs are 0.25 for EUK station and 0.85 for SUM and BRW, see Fig. 3.3.2.

For stations located in Asia (RYO, YON, MNM) both runs slightly overestimate O_3 mixing ratios with MNMBs between 14% and 19% for the o-suite and between 15% and 21% for the control run, see Fig 3.3.6. Correlation coefficients range between 0.33 and 0.94.

For MBO and NWR USA stations, the observed ozone mixing ratios are overestimated by the o-suite by 25% and 23% respectively. The control run MNMBs are slightly lower (MNMB$_{control}$=20% at MBO
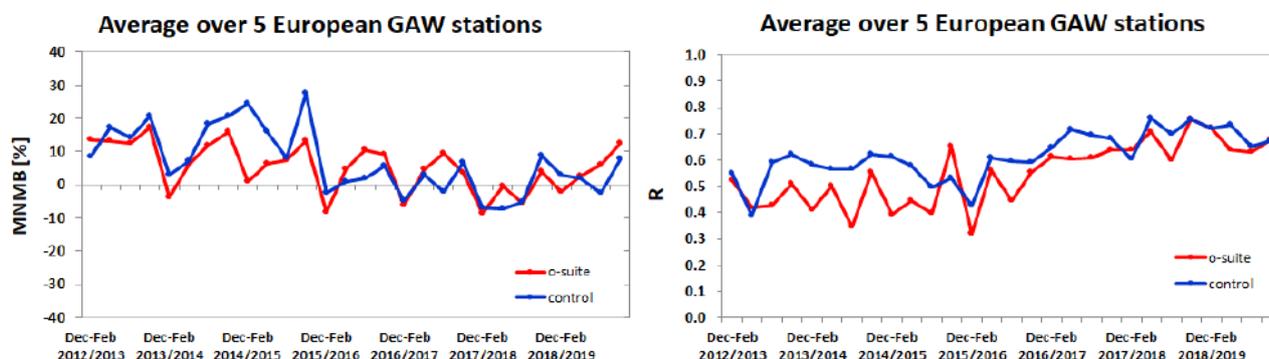
Figure 3.3.3: Long term (Dec. 2012 – November 2019) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).
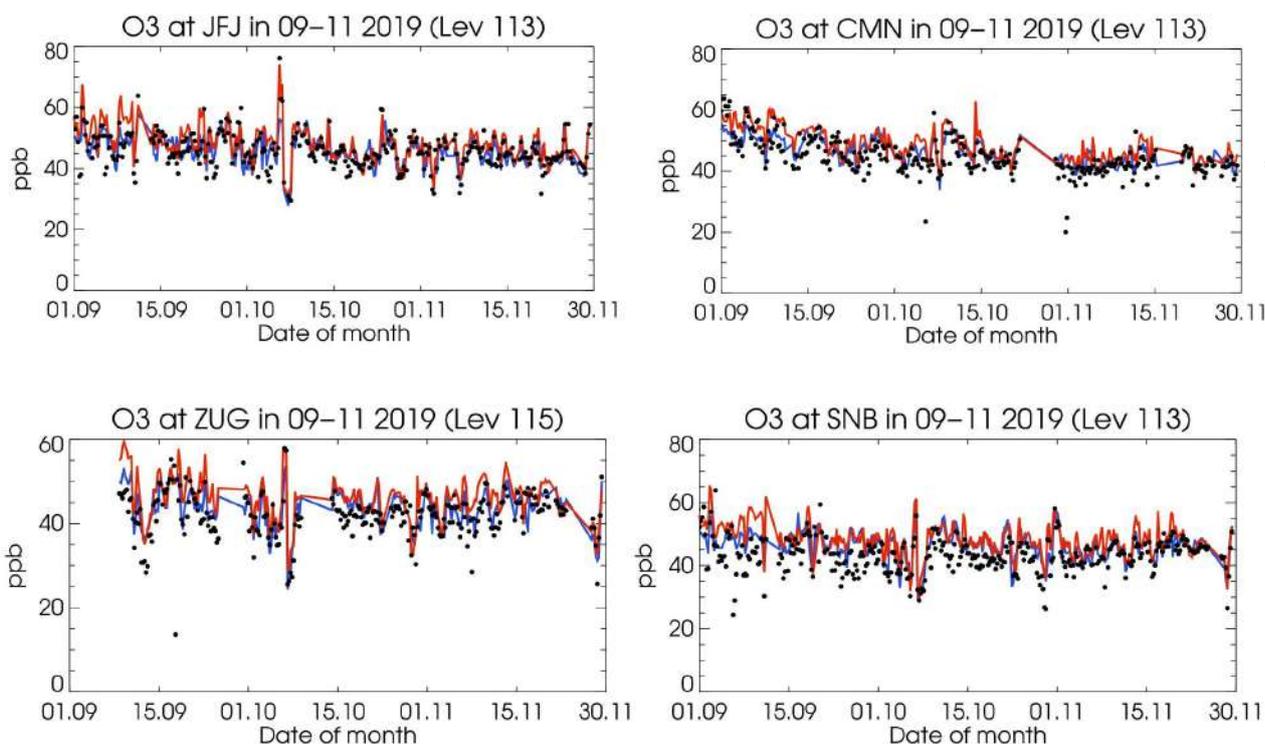

Figure 3.3.4: Time series for the o-suite (red) and control (blue) compared to GAW observations for Jungfraujoch (46.55°N, 7.99°E) and Cimone (44.18°N, 10.70°E) (upper panel). Zugspitze (47.4°N, 10.9°E) and Sonnblick (47.05°N, 12.96°E) (lower panel).

and 19% at NWR). Correlations between o-suite and observations are between 0.62 (at MBO) and 0.49 (at NWR) for the o-suite and between 0.65 (at MBO) and 0.49 (at NWR) for the control run.

The observed ozone mixing ratios are overestimated by both runs over Mauna Loa (MLO) station in the Tropics (MNMB$_{o-suite}$≈13%, MNMB$_{control}$≈15%), see Fig. 3.3.7. Correlations between simulated and observed surface ozone are high for both the o-suite and the control run (r>0.8).
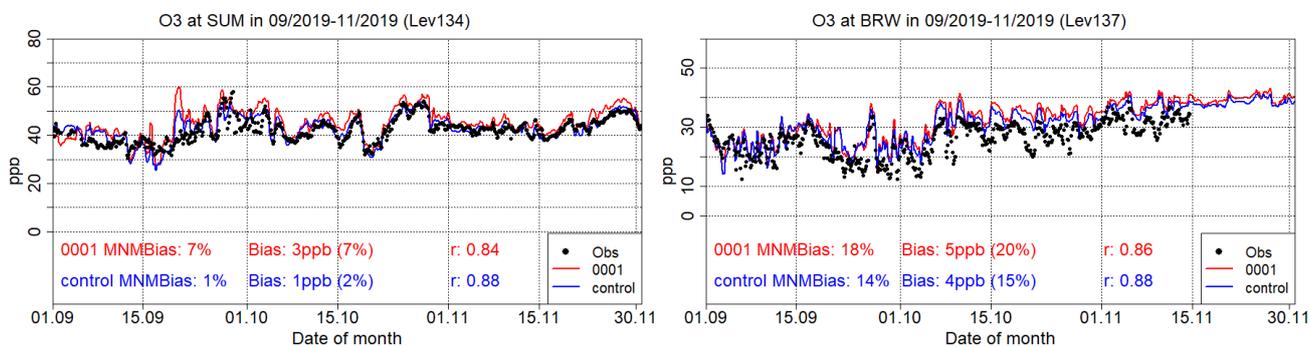
Figure 3.3.5: Time series for the o-suite (red) and control (blue) compared to ESRL observations at Summit, Greenland station (72.57°N, 38.48°W, left) and at Point Barrow, Alaska station (71.32°N, 156.51°W, right)
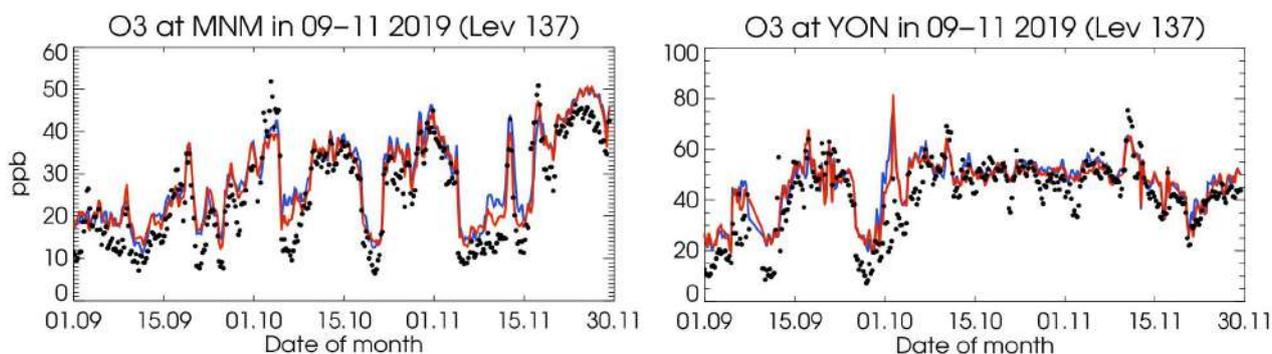

Figure 3.3.6: Time series for the o-suite (red) and control (blue) compared to GAW observations for Minamitorishima (24.29°N, 153.98°E) and Yonagunijima (24.47°N, 123.02°E).
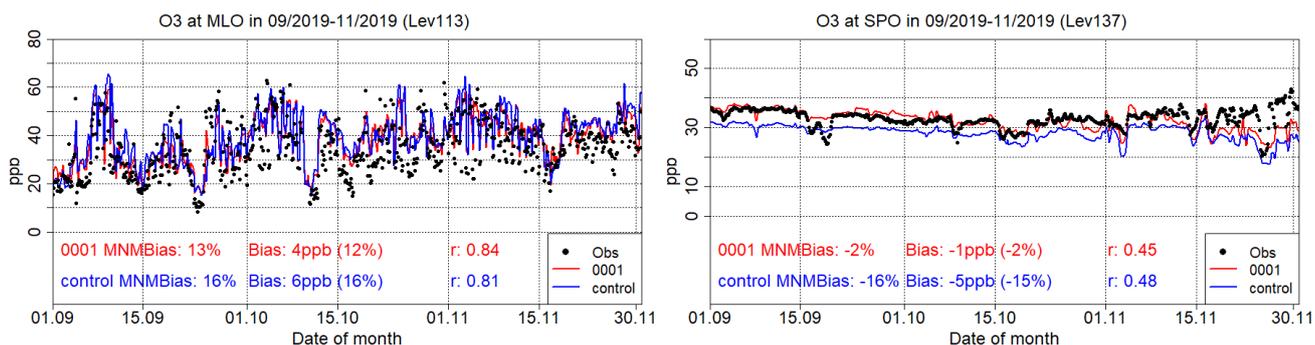

Figure 3.3.7: Time series for the o-suite (red) and control (blue) compared to ESRL observations (black dots) at Mauna Loa, Hawaii station (19.54°N, 155.58°W) and at South Pole, Antarctica station (90.00°S, 24.80°W).

At CVO station, the model corresponds very well to the observations with almost 0 MNMB and correlation of 0.86 (Fig. 3.3.9).

The O_3 mixing ratios of the southern hemispheric stations (CPT, USH) show MNMBs between 6 and -5% for the o-suite. The control run shows larger underestimations for USH up to -20%, see Fig 3.3.8. At Lauder (LDR) station in New Zealand the o-suite overestimates O_3 mixing ratios by 19% and the control run by 6%. Correlations between simulated and observed surface ozone values for the o-suite and the control run are 0.73 and 0.78 respectively.
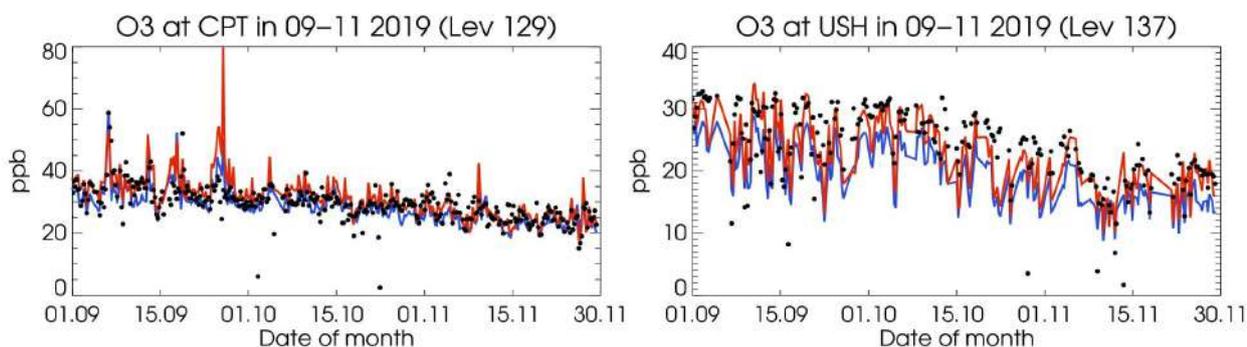
Figure 3.3.8: Time series for the o-suite (red) and control (blue) compared to GAW observations (black dots) at Cape Point (34.55°S, 18.48°W) and GAW observations at Ushuaia (5.85°S, 68.32°W).
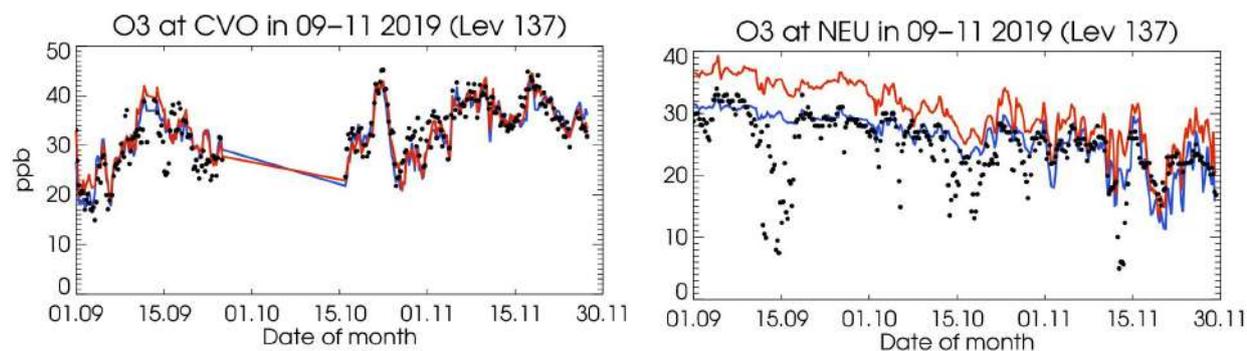

Figure 3.3.9: Time series for the o-suite (red) and control (blue) compared to GAW observations (black dots) at Cape Verde (16.85°N, 24.87°W) and GAW observations at Neumayer (70.65°S, 8.25°W).

At Arrival Height (ARH) station in Antarctica the MNMB is 15% for the o-suite and almost zero for the control run. Correlation coefficients are very poor for both runs (r≈0). Finally, for South Pole station in Antarctica (SPO), the MNMB is -2% for the o-suite and -15% for the control run. Correlation coefficients 0.45 for the o-suite and 0.48 for the control run.

For Neumayer station (NEU) the MNMB is 22% for the o-suite and 7.5% for the control run. Correlation coefficients are 0.57 for the o-suite and control run, Fig. 3.3.9.

3.4 Validation with AirBase observations in Mediterranean

The surface ozone validation analysis over the Mediterranean is based on an evaluation against station observations from the Airbase Network (http://acm.eionet.europa.eu/databases/airbase/). In addition, 1 station from the Department of Labour Inspection - Ministry of Labour and Social Insurance, of Cyprus (http://www.airquality.dli.mlsi.gov.cy/) is used in the validation analysis. For the validation analysis, stations in the Mediterranean located within about 100 km from the shoreline of the Mediterranean shore are used. Table 3.4.1 shows the names, coordinates, elevation and the MNMBs and correlations obtained with the 2 forecast runs (o-suite and control). It indicates that the variance explained by each station of both the o-suite and control is high and correlations are highly significant over Western, Central and Eastern Mediterranean. It should be

Table 3.4.1: Coordinates, elevation, corresponding model level (level 137 is the surface level), as well as validation scores (MNMBs and correlations for the period SON 2019) obtained with the 2 forecast runs (o-suite and control), for each one of the selected Mediterranean stations. MNMBs and correlations with blue denote stations where control run performs better while with red are denoted stations where o-suite performs better.

Station Name	Stat_ID	Lon	Lat	Alt (m)	Level	Distance from the shore (km)	MNMB		Cor. Coef	
							o-suite	control	o-suite	control
Al Cornocales	ES1648A	-5.66	36.23	189	133	16	25.4	20.2	0.75	0.74
Caravaka	ES1882A	-1.87	38.12	1	137	73	14.1	8.6	0.79	0.73
Zarra	ES0012R	-1.10	39.08	885	130	70	8.2	3.3	0.90	0.87
VIillar Del Arzobispo	ES1671A	-0.83	39.71	430	137	48	2.0	-2.3	0.36	0.40
Cirat	ES1689A	-0.47	40.05	466	137	37	31.5	27.5	0.44	0.45
Bujaraloz	ES1400A	-0.15	41.51	327	137	60	17.3	12.8	0.65	0.62
Morella	ES1441A	-0.09	40.64	1150	128	51	NA	NA	NA	NA
Bc-La Senia	ES1754A	0.29	40.64	428	137	21	-13.8	-18.2	0.72	0.71
Ay-Gandesa	ES1379A	0.44	41.06	368	136	15	20.5	15.8	0.67	0.66
Ak-Pardines	ES1310A	2.21	42.31	1226	135	81	26.6	21.5	0.64	0.61
Hospital Joan March	ES1827A	2.69	39.68	172	133	3	17.8	13.1	0.58	0.56
Al-Agullana	ES1201A	2.84	42.39	214	137	25	5.6	0.7	0.75	0.73
Av-Begur	ES1311A	3.21	41.96	200	132	9	29.5	24.9	0.83	0.81
Plan Aups/Ste Baume	FR03027	5.73	43.34	675	124	21	21.6	17.0	0.76	0.76
Montemonaco	IT1842A	13.34	42.90	1000	127	46	15.9	11.5	0.79	0.84
Gharb	MT00007	14.20	36.07	114	132	31	1.0	-2.1	0.71	0.65
Aliartos	GR0001R	23.11	38.37	110	59	18	NA	NA	NA	NA
NEO	-	21.67	37.00	50	60	2	NA	NA	NA	NA
Finokalia	GR0002R	25.67	35.32	250	57	4	5.2	2.9	0.91	0.86
Agia Marina	CY0002R	33.06	35.04	532	55	14	1.3	0.8	0.78	0.82

noted that the control run reproduces slightly better than the o-suite run the surface ozone day to day variability over almost all the Mediterranean stations (see Table 3.4.1, exception is the Agia marina Cyprus station).

In terms of biases, o-suite mostly overestimate surface ozone values and its MNMBs vary between -14% and +30% depending on the stations over the Mediterranean shore of Spain (average MNMB for the 13 Spain Mediterranean station is 15%). The Control MNMBs are on average 4% lower than o-suite MNMBs and are closer to zero compared to o-suite MNMB. Over the stations Plan Aups/Ste Baume in France and Montemonaco in Italy the o-suite overestimate surface ozone concentrations by 21% and 16% respectively. Again, the Control MNMBs are lower 4.5% than o-suite MNMBs and are closer to zero compared to o-suite MNMB. Over Gharb station in Malta the o-suite overestimate surface ozone values by 1% and the control run underestimate it by -2%. Over Finokalia station in Crete the o-suite overestimate surface ozone by 5% while the control run overestimate it by 3%. Finally, over Agia Marina in Cyprus both runs overestimate surface ozone values by 1%.

The spatial distribution of MNMBs and the correlation coefficients of the o-suite over the Mediterranean are shown in 3.4.2, where it is evident that correlations over the entire Mediterranean from Gibraltar to Cyprus are highly significant. It is also evident that the CAMS NRT runs have a better performance over Central and eastern Mediterranean compared to the Mediterranean shore of Spain in terms of biases.
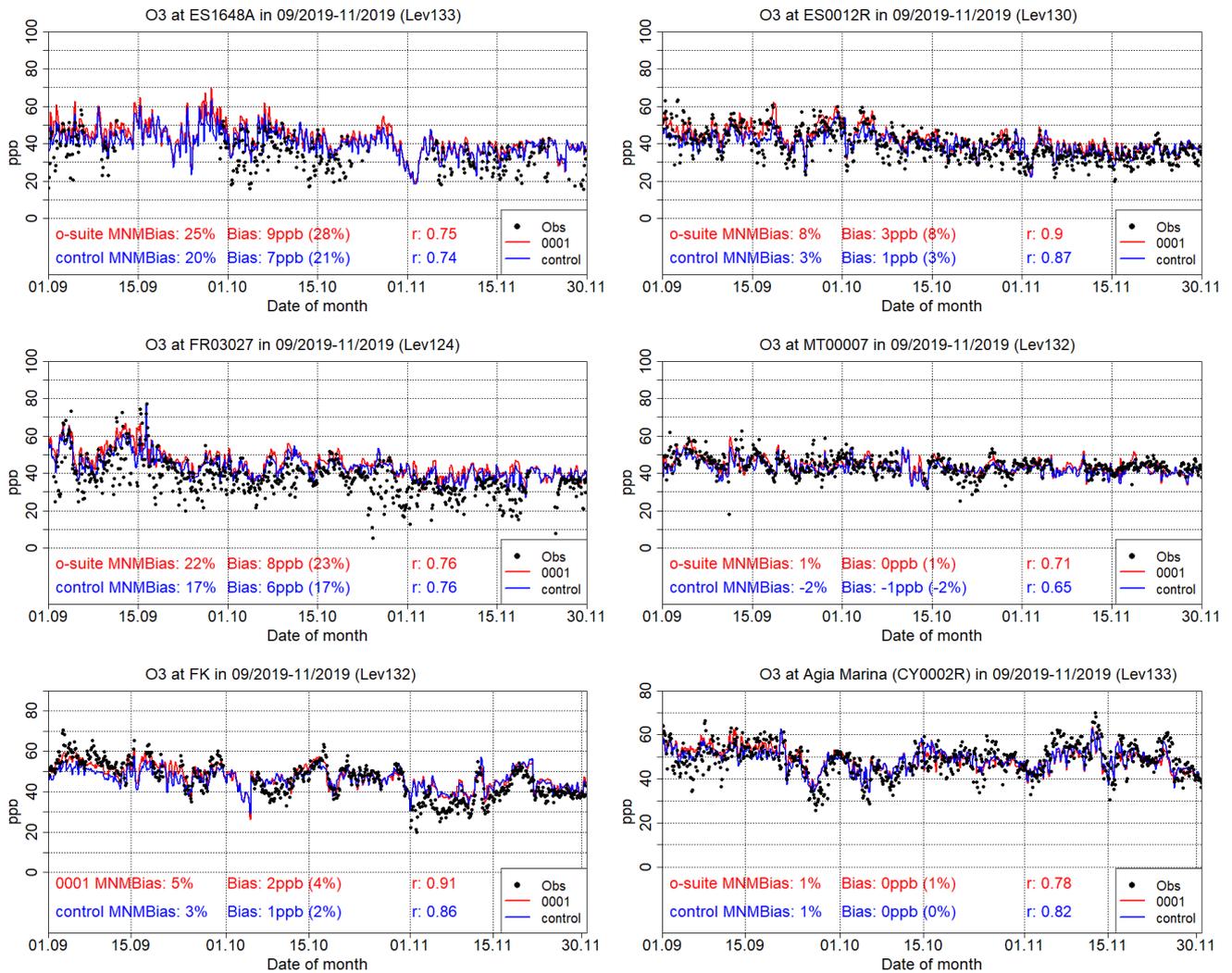
Figure 3.4.1: Time series for the o-suite (red) and Control (blue) compared to Airbase observations at Al Cornocales, Spain station (36.23°N, 5.66 °W, top left), at Zarra, Spain station (39.08°N, 1.10°W, top right), at Plan Aups/Ste Baume, France station (43.34°N, 5.73°E, center left), at Gharb, Malta station (36.07°N, 14.20°E, center right at Finokalia, Crete Greece station (35.32°N, 25.67°E, bottom left) and compared to observations provided by the Department of Labour Inspection - Ministry of Labour and Social Insurance of Cyprus) at Agia Marina, Cyprus station (35.04°N, 33.06 °E, low right).
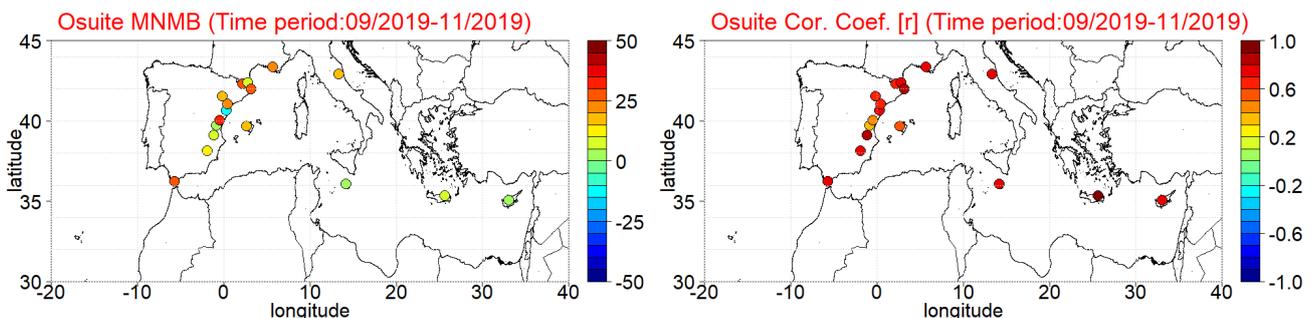
Figure 3.4.2: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 September 2019 to 30 November 2019.
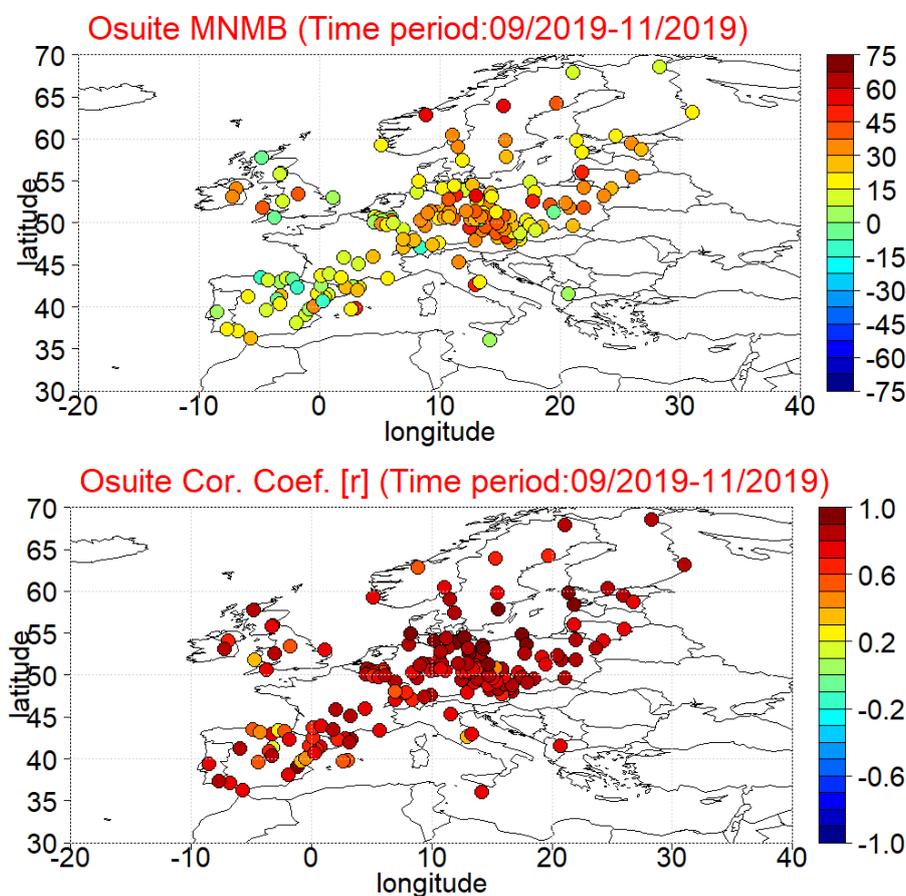
Figure 3.5.1: Spatial distribution of MNMB in % (left) and correlation coefficient (right) of the o-suite run compared to observational data during the period from 1 September 2019 to 30 November 2019.

3.5 Validation with AirBase observations over Europe

The surface ozone validation analysis over Europe is based on an evaluation against Background rural Classes 1-2 O_3 July-Peuch classification station observations from Airbase Network (http://acm.eionet.europa.eu/databases/airbase/). The spatial distribution of MNMBs and the correlation coefficients of the o-suite over Europe are shown in 3.5.1, where it is evident that correlations over most European AirBase stations (with a very few exceptions) are highly significant (0.5<r<0.95). It is also evident that the CAMS NRT runs reproduce well the surface ozone mean concentrations over the Iberian Peninsula and France (depending on the station MNMBs vary from -20% to +20%). It is also evident that over the rest of Central and Northern Europe the o-suite mostly overestimates surface ozone values up to 50%. The above-mentioned findings concerning CAMS NRT runs biases and correlations are also observed in individual time series at selected stations plotted in Figure 3.5.2. From this time series and the plotted validation metrics is also evident that control run surface ozone concentrations are 1-2 ppb (4%-8%) lower than the o-suite values resulting in most cases a closer to zero bias.
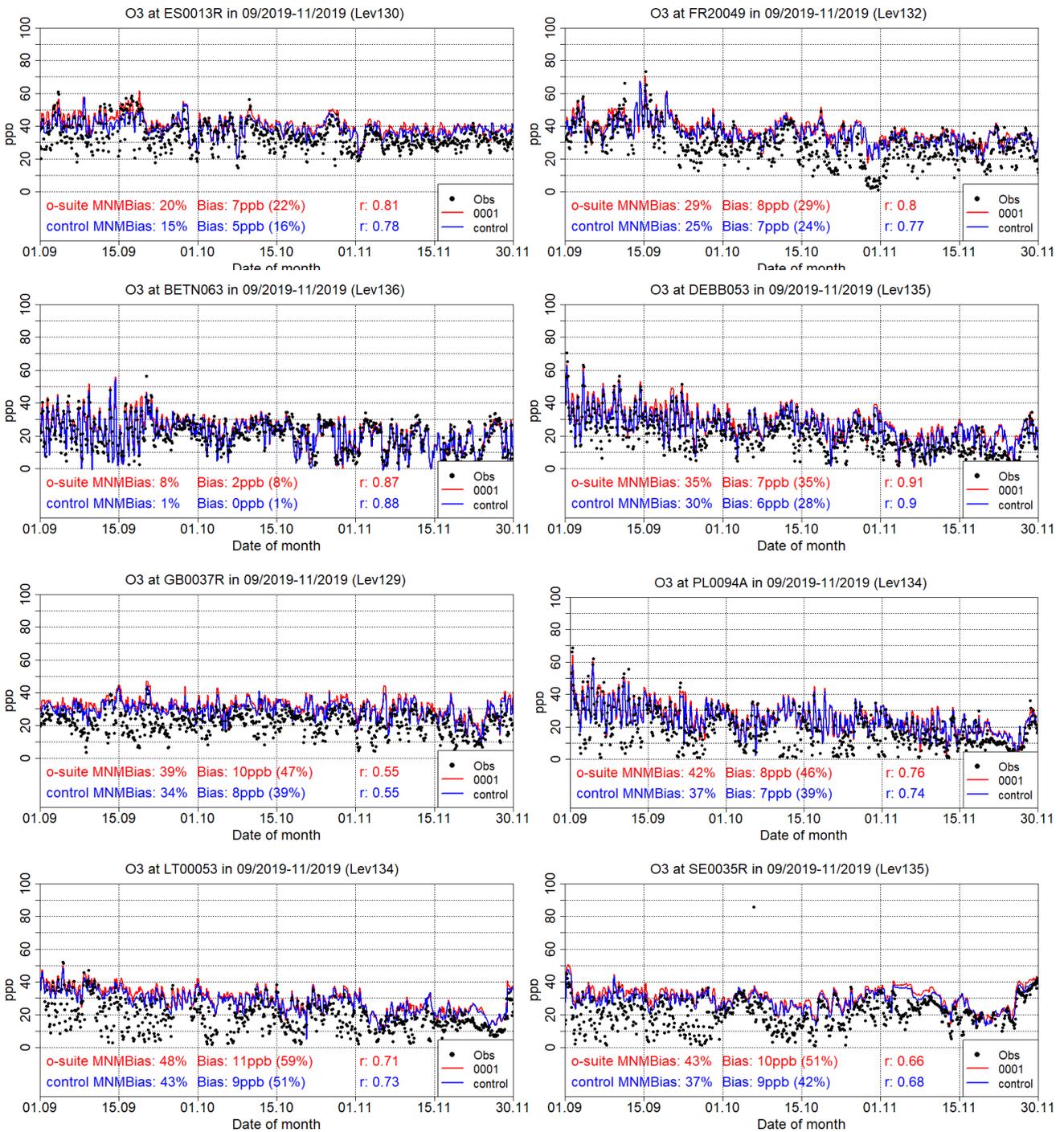
Figure 3.5.2: Time series for the o-suite (red) and Control (blue) compared to Airbase observations at Al Penausende, Spain station (41.24°N, 5.90 °W, 1st row left), at Haut Beaujolais, France station (45.96°N, 4.47°E, 1st row right), at Corroy L.G., Belgium Station (50.67°N, 4.67°E, 2nd row left), at Hasenholz, Germany (52.56°N, 14.02°E, 2nd row rigth), at Ladybower, Great Britain station (53.40°N, 1.75°W, 3nd row left), at LdGajewWIOSAGajew, Poland station (52.14°N, 19.23°E 3nd row rigth), at Zemaitija, Lithuania station (56.01°N, 21.89°E, 4nd row left) and at Vindeln, Sweden station (64.25°N, 19.77°E, 4nd row right).
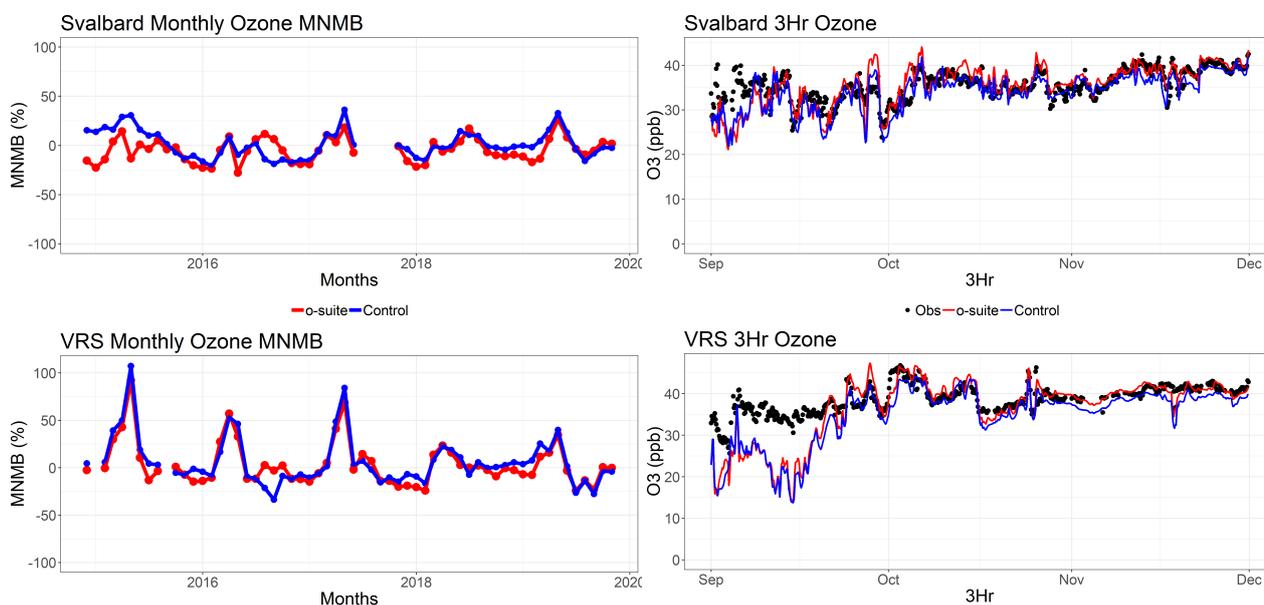
Figure 3.6.1: Time series for o-suite (red) and Control (blue) compared to observations (black dots) at, Svalbard (top row), and the Villum Research Station, Station Nord, Greenland (bottom row) MNMB for the full period (left) and concentrations for September-November 2019 (right).

3.6 Validation with IASOA surface observations

Model results were compared to surface O_3 observations from the Villum Research Station, Station Nord in north Greenland (81.6°N 16.7°W), and Zeppelin Mountain, Svalbard (78.9°N 11.9°E) from the IASOA network (Fig. 3.6.1).

The data from Svalbard and VRS are covering the period from December 2014 to November 2019. The model simulations do not capture ozone depletion events in March – June in 2015 – 2019 at any of the sites. These events are related to halogen chemistry reactions that are not represented in the model simulations. The simulations are on average in good agreement with the observations apart from the spring depletion events.

For the period September to November 2019 the measurements are not quality controlled. The simulated O_3 concentrations at the two sites are in well agreement with the measurements resulting in low negative bias of 0% - 7% for the o-suite and 4% - 12% for the control run (Table 3.6.1). The bias at VRS is mainly due to underestimation in the beginning of September, whereas the model is in excellent agreement for the rest of the period. High correlation coefficients are found at both sites with slightly better performance for the control run: r = 0.69 – 0.81 for the o-suite compared to r = 0.75 – 0.81 for the control run.

Table 3.6.1. Modified Normalised Mean Bias (MNMB) and correlation coefficient (r) of the o-suite and the Control simulations for the sites Svalbard, and Villum Research Station (VRS) for the period Sep-Nov 2019.

		MNMB	R
Svalbard	o-suite	-0.00	0.69
	control	-0.04	0.75
VRS	o-suite	-0.07	0.81
	control	-0.12	0.81

3.7 Validation with IASI data

Ozone total columns from the o-suite and control run are compared with IASI Metop-A version v20151001 daytime only satellite observations (Clerbaux et al., 2009). For the comparison with the IASI data, the vertically integrated model O_3 data were transformed using IASI averaging kernels (Rodgers, 2000).

The global distribution of the O_3 total column obtained from IASI, as well as the relative difference between the model runs and IASI, are shown in Fig. 3.7.1 for September 2019. Satellite data shows high O_3 over the southern mid- and high-latitude, especially over the Indian and Pacific Oceans and low values over the ocean south of Patagonia and equatorial areas. The o-suite run captures both, high and low O_3 values relatively well and is in good agreement with the observations, showing MNMBs within 5%. The control run is mainly positively biased (up to 20% over the high northern latitudes). Slightly negative biases appear over the biomass burning area in Africa (within 10%). The forecast day 4 is almost similar to the forecast day 0. Note, that the IASI sensitivity is lowest over the cold surfaces of Antarctica and Greenland (especially during March-April-May season) where IASI O_3 values are positively biased by up to 20%.
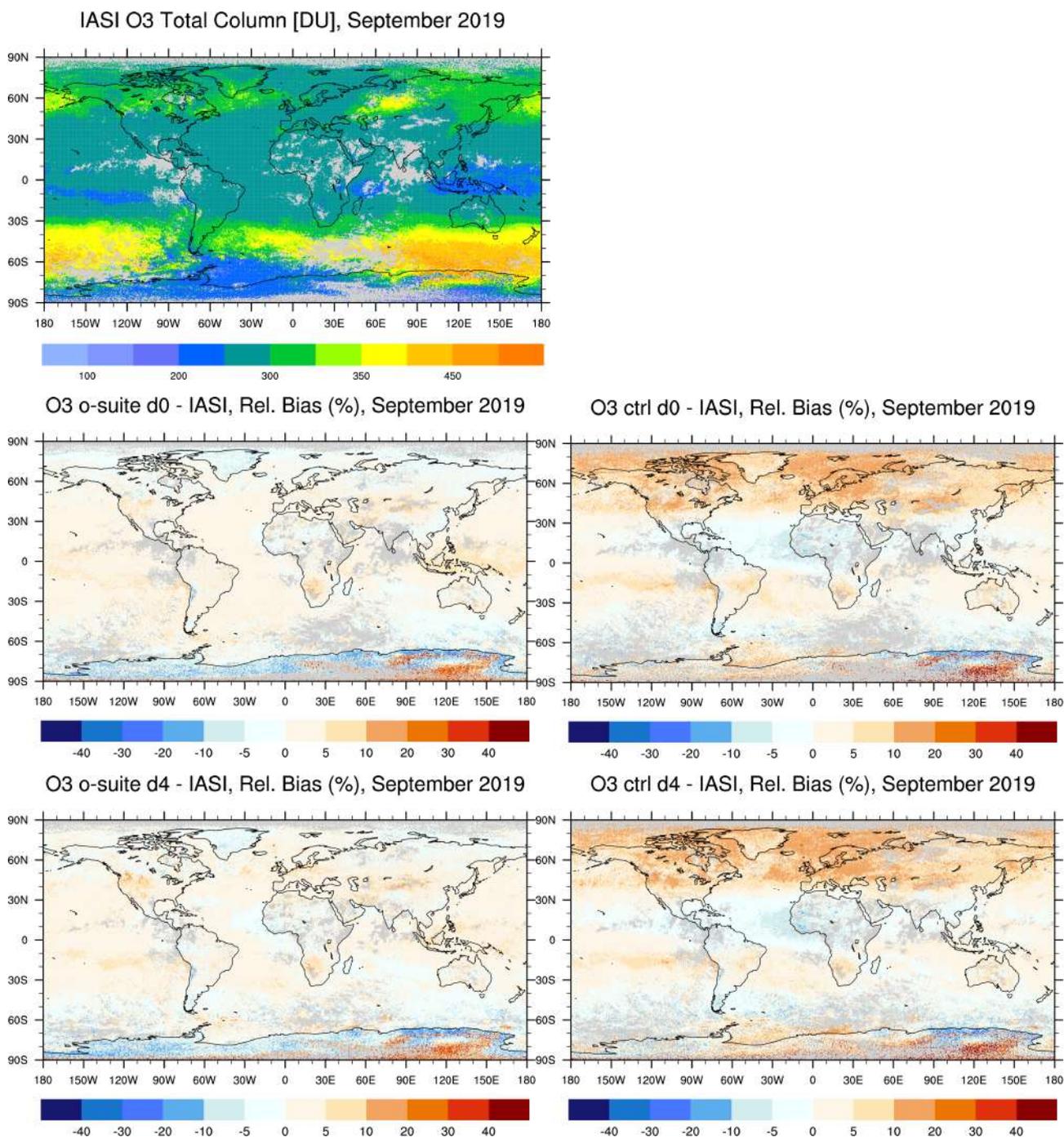
Figure 3.7.1: O_3 total column for IASI satellite observations (top) and relative difference between the model runs and IASI for September 2019: o-suite day 0 and day 4 (left), control run day 0 and day 4 (right). Grey colour indicates missing values.

4. Carbon monoxide

4.1 Validation with Global Atmosphere Watch (GAW) Surface Observations

For the Near-Real-Time (NRT) validation, 10 GAW stations have delivered CO surface mixing ratios in NRT and data is compared to model results as described in Eskes et al. (2019) and is used for CAMS model evaluation for September 2019 to November 2019. The latest validation results can be found on the CAMS website, see section 1.

For stations in the Northern Hemisphere, both runs mostly show slightly negative MNMBs for stations in Europe. The control run has a slightly larger negative offset than the o-suite (Fig. 4.1.1).

A comparison of the seasonal-mean MNMB over Europe (Fig. 4.1.2) from December 2012 to present shows a slowly improving MNMB from about -20% in 2013 to about -10% for more recent periods. Temporal correlation remains relatively constant at r=0.6 on average, except for the quarter JJA in 2018, where the correlation of the control run drops to 0.24.
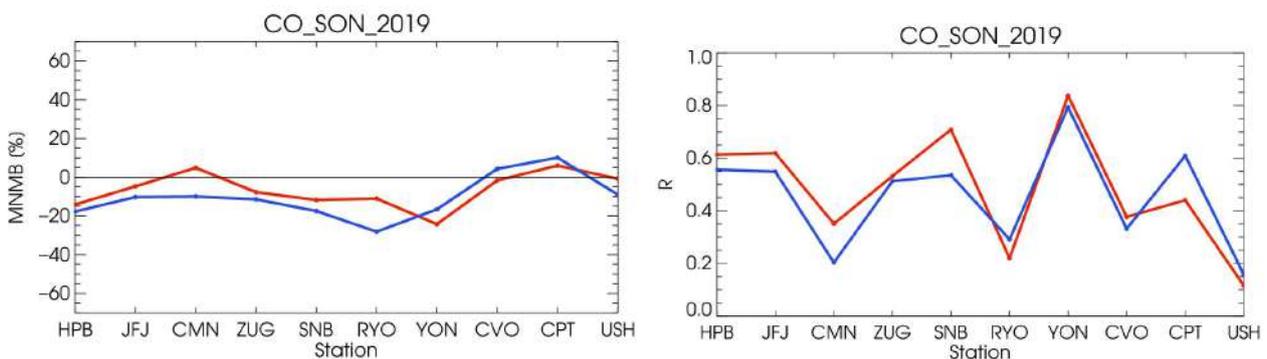
Figure 4.1.1: Modified normalized mean bias in % (left) and correlation coefficient (bottom right) of the NRT model runs compared to observational GAW data in the period September 2019 to November 2019 (o-suite: solid red, and control: blue).
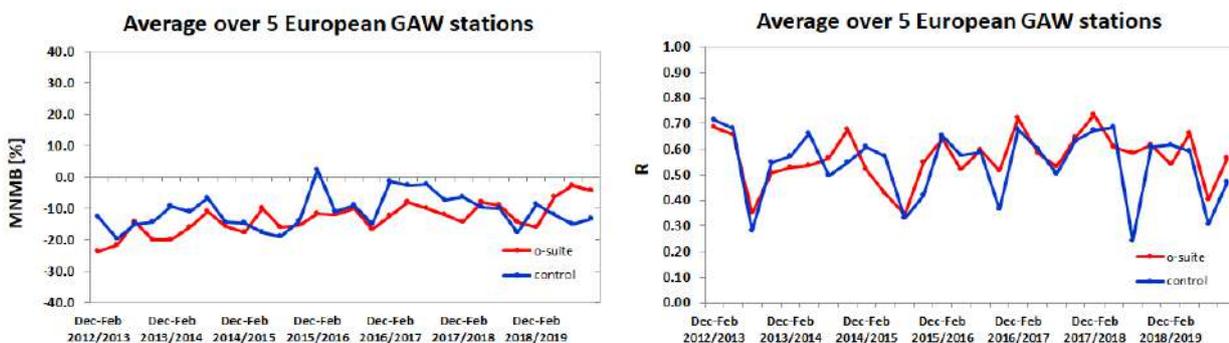
Figure 4.1.2: Long term (Dec. 2012 – November 2019) evolution of seasonal mean MNMB (left) and correlation (right), as averaged over 5 GAW stations in Europe, for o-suite (red) and control (blue).
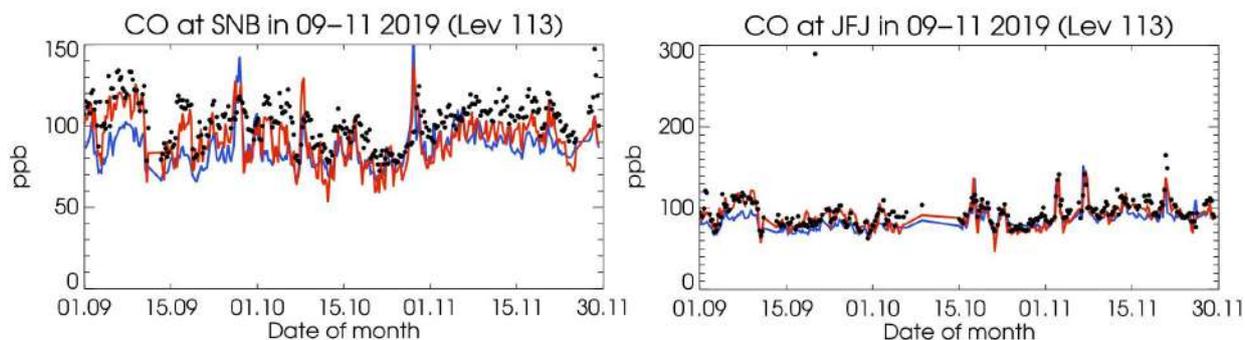
Figure 4.1.3: Time series for the o-suite (red) and control (blue) compared to GAW observations at Sonnblick (47.05°N, 12.96°E) and Jungfraujoch (46.55°N, 7.99°E).
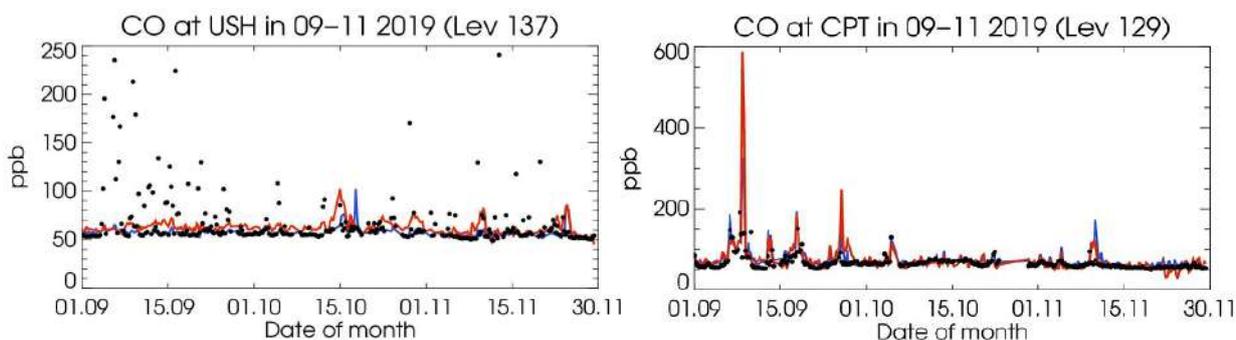

Figure 4.1.4: Time series for the o-suite (red) and control (blue) compared to GAW observations at Ushuaia (54.85°S, -68.32°W) and Cape Point (34.35°S, 18.5°E).
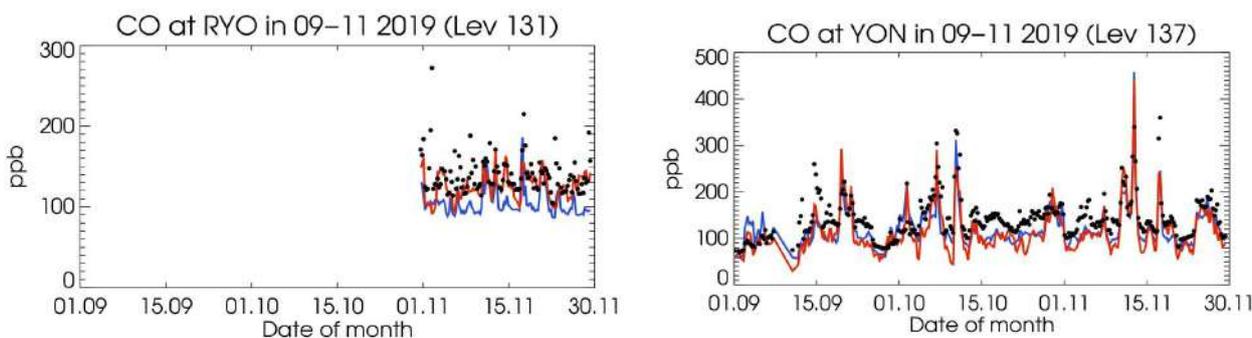

Figure 4.1.5: Time series for the o-suite (red) and control (blue) compared to GAW observations for Ryori (39.03°N, 141.82°E) and Yonagunijima (24.47°N, 132.02°E)

For European stations (Fig. 4.1.3), the o-suite shows an underestimation of observed CO mixing ratios, with MNMBs between 4% and -14%. The control shows slightly larger negative MNMBs between -4% and -17%. Correlation coefficients are between 0.35 and 0.7 for the o-suite and between 0.2 and 0.55 for the control run.

For the two stations in the Southern Mid-latitudes (CPT and USH) MNMBs are low (Fig.4.1.4.) for o-suite and control. For stations in Asia (RYO, YON) both runs underestimate CO mixing ratios by up to -24% (up to -28% for the control run), see Fig. 4.1.5. Correlation coefficients are slightly better for the o-suite (RYO 0.29, YON 0.83) than for the control run (RYO 0.29, YON 0.79).
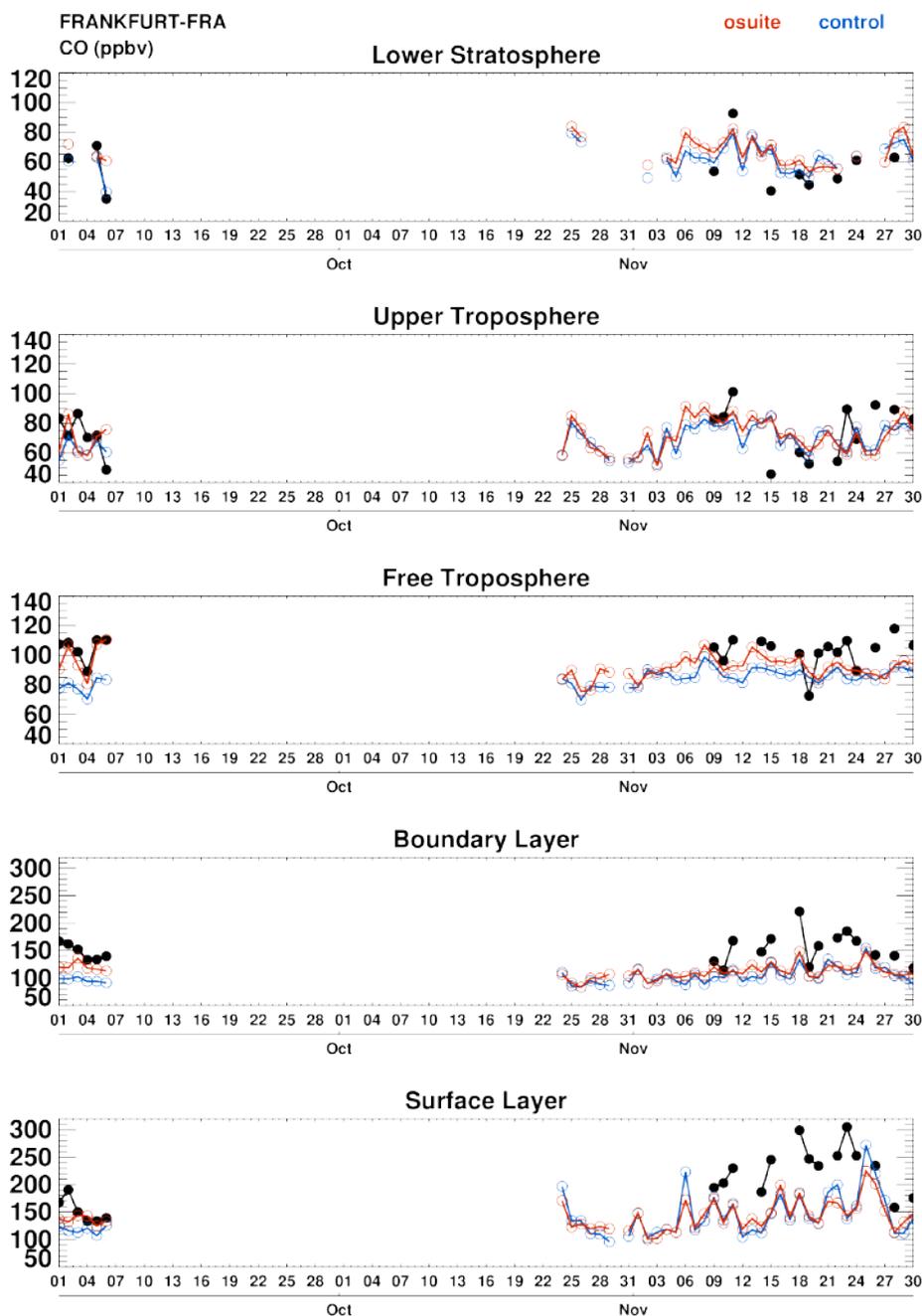
Figure 4.2.1 Time series of daily mean CO over Frankfurt during SON 2019 for 5 layers: Surface Layer, Boundary Layer, Free Troposphere, Upper Troposphere and Lower Stratosphere. The o-suite is shown in red and associated control run in blue.

4.2 Validation with IAGOS Data

Very poor sampling is available for CO at Frankfurt during SON 2019 (Fig. 4.2.1 and 4.2.2). CO is mostly underestimated by both o-suite and associated control run and the largest bias is generally found in the lowest layers. While the performances of the two runs are more similar in the lowest layers, the o-suite performs slightly better than the control run in the free troposphere and UTLS.
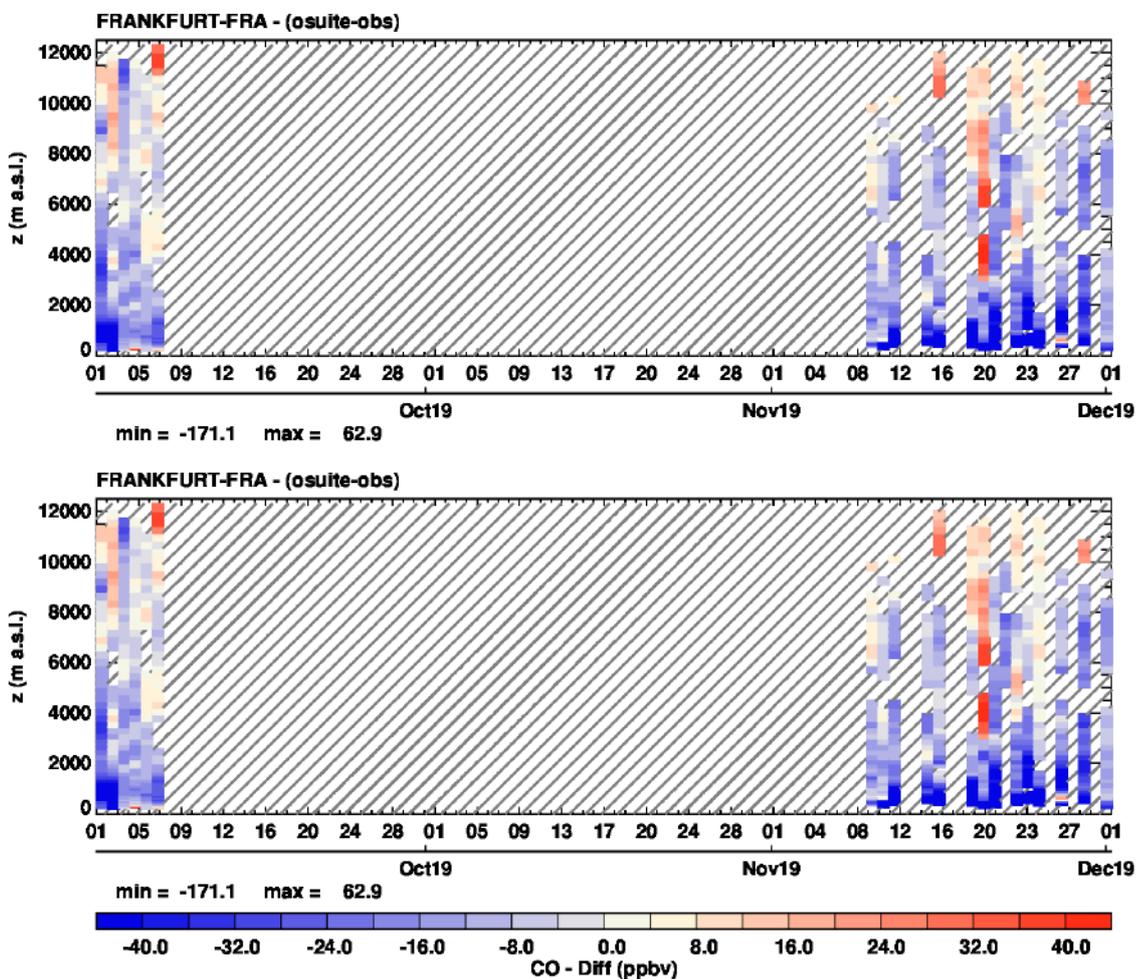
Figure 4.2.2 Time series of the absolute differences (model - observations) in daily profiles for CO over Frankfurt during SON 2019. Two upper panels correspond to o-suite (before and after upgrade) and two lower panels to control run (before and after upgrade).

High CO values around 300 ppbv are observed near the surface on several days in November as shown in the individual profiles of Fig. 4.2.3 (1118, 1119, 1123 and 1124). For these profiles, the two runs behave similarly and surface values are underestimated by about 30%. In the boundary layer the models underestimate CO as well but the bias is slightly smaller. In the free troposphere and UTLS the bias is smaller than in the low troposphere and according to all individual profiles the performance from the o-suite is in general better than that of the control. In the profile of 1119, CO values are particularly low in the free troposphere with values below 50 ppbv, and this might be related to the stratospheric intrusion observed in the corresponding ozone profile (see ozone section). In this case, CO is overestimated by both models in the range of altitude between 3000 m and 9000m.
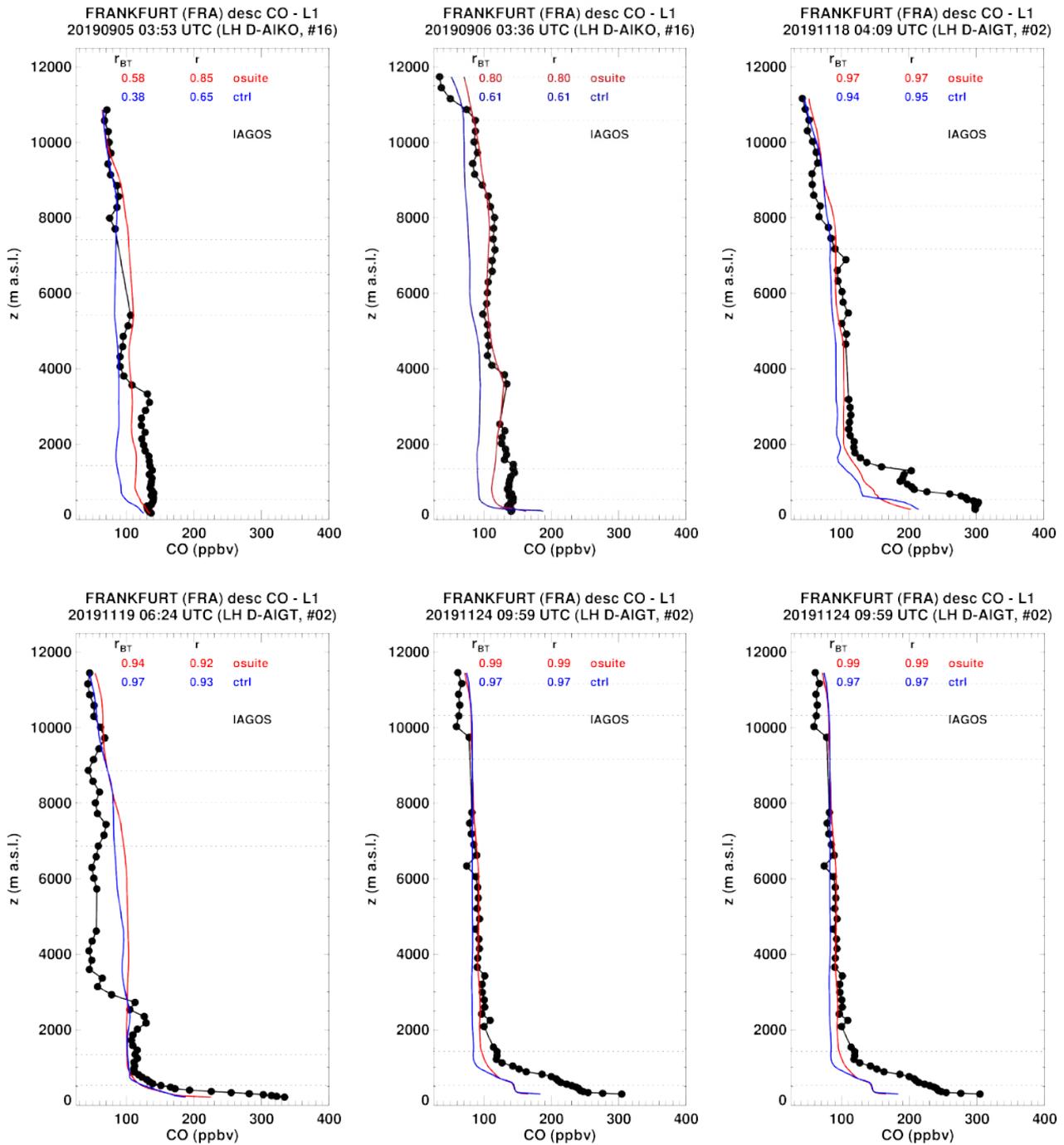
Figure 4.2.3 Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Europe during SON 2019.
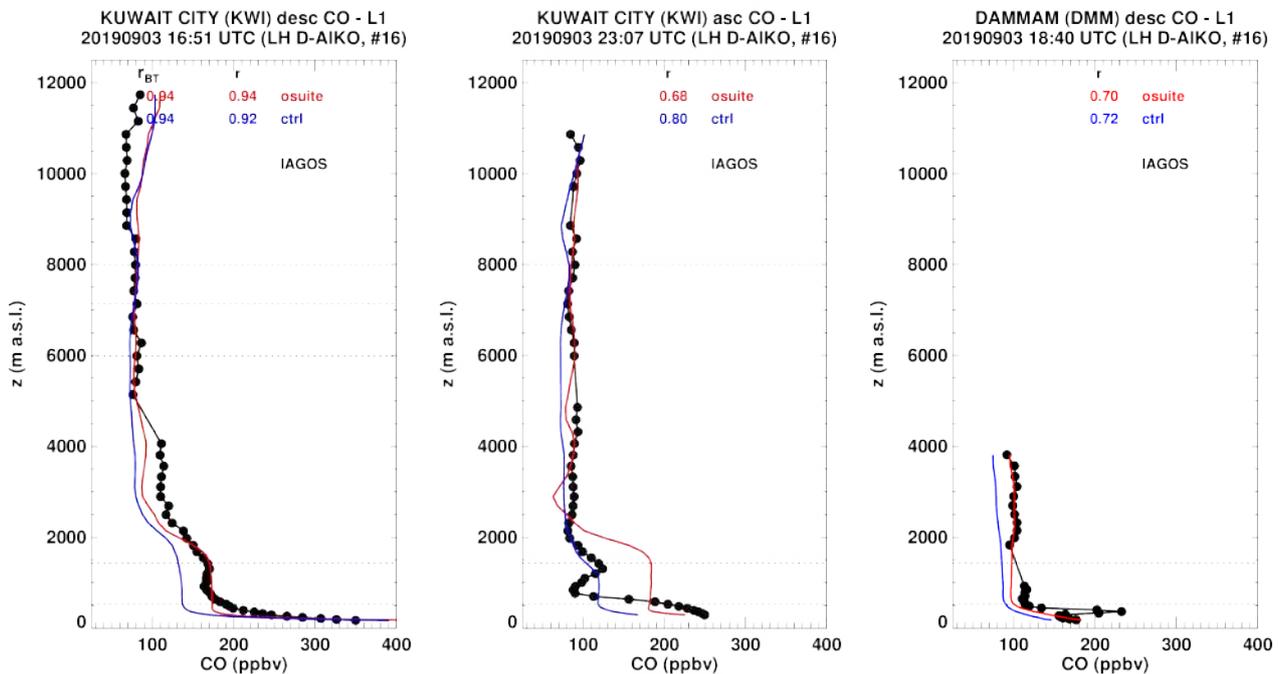
Figure 4.2.4: Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over the Middle East during SON 2019.

Middle East

A few profiles are available over the Middle East at the airports of Kuwait City and Dammam (Fig. 4.2.4). In the first profile at Kuwait city on 3 September, high mixing ratio of CO reaching more than 400 ppbv is observed in the surface layer and 180 ppbv in the boundary layer. CO values are well reproduced by both models. However, in the boundary layer the models differ and the o-suite performs better than the control run. On the second profile at Kuwait city on the same day, the o-suite performs better than control run near the surface while control run present better performances in the boundary layer. In the UTLS, the behaviour of the two runs is similar and both agree well with observations.
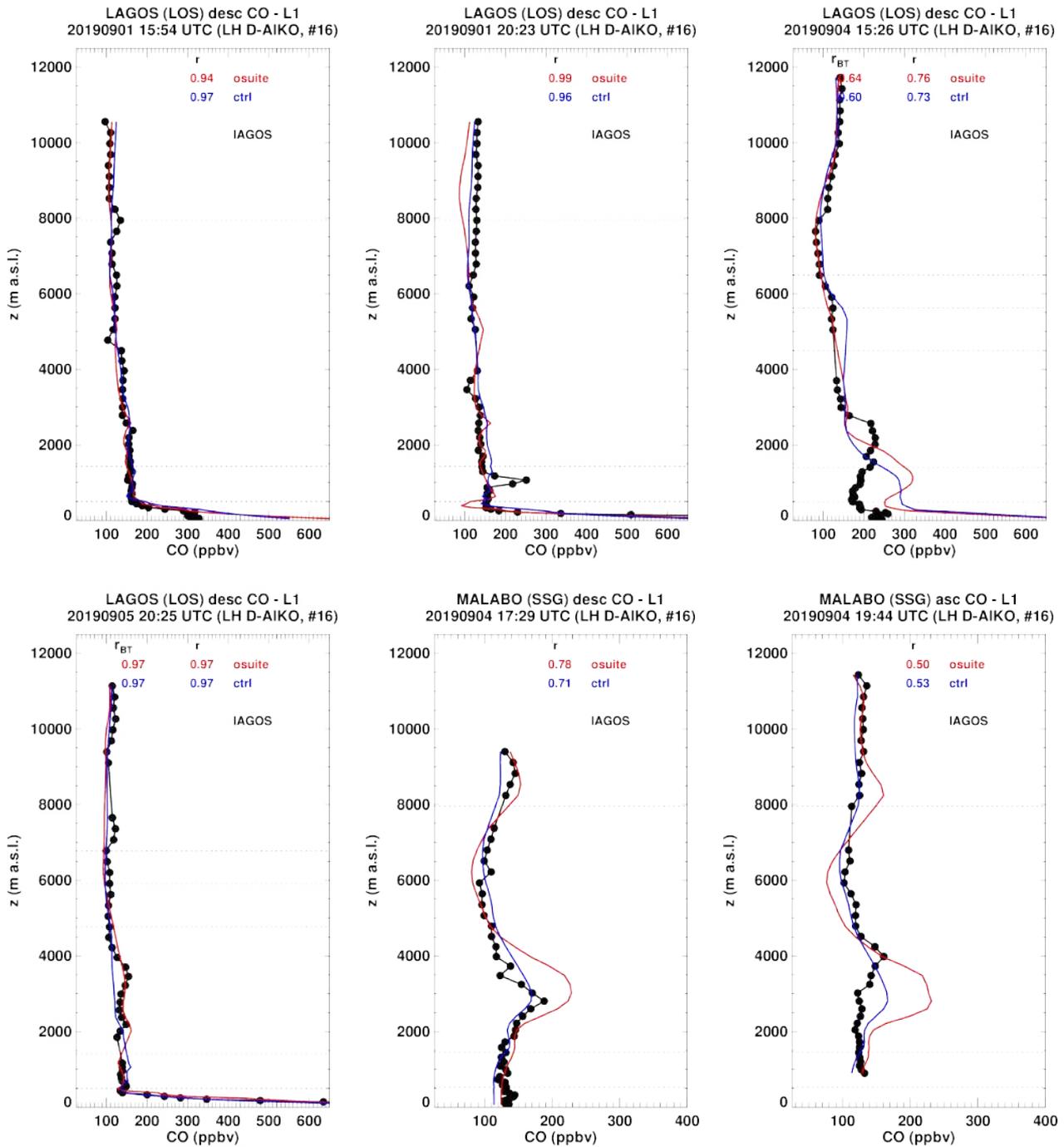
Figure 4.2.5: Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over Western Africa during SON 2019.
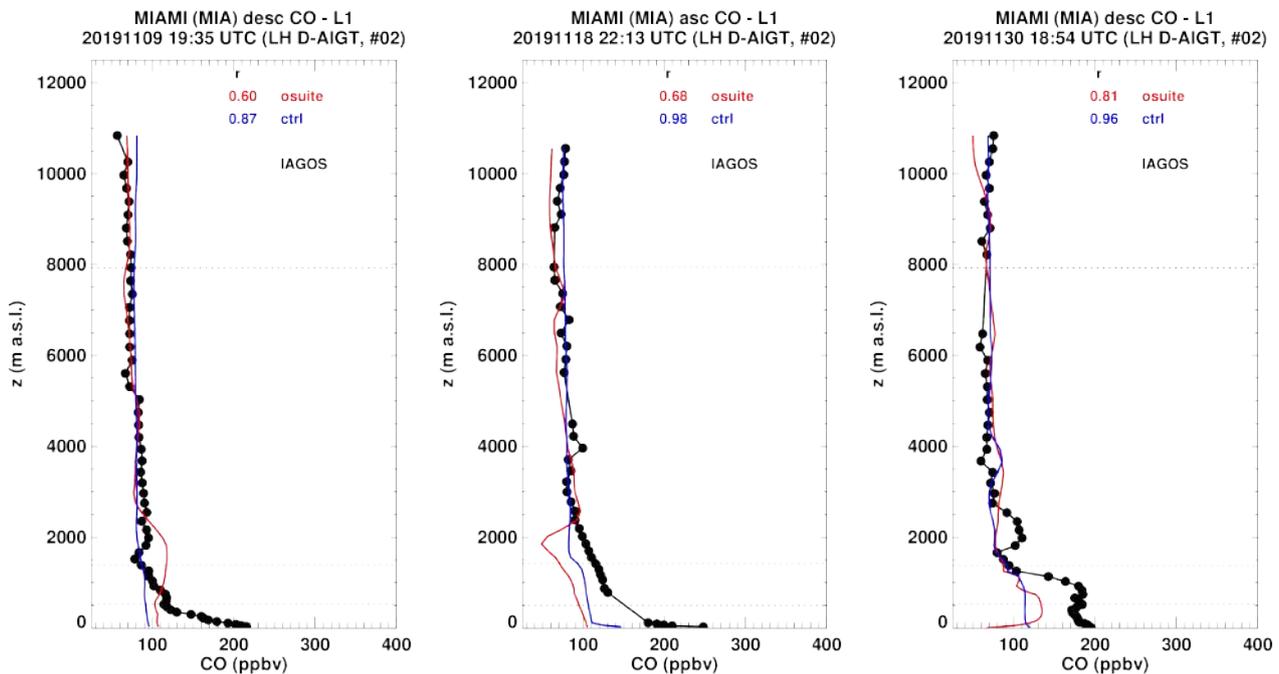
Figure 4.2.6: Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over North America during SON 2019.

Africa

Over Africa, CO profiles are available at Lagos and Malabo both located in the gulf of Guinea (Fig. 4.2.5). High mixing ratios of CO observed near the surface at Lagos are well represented by the two runs (0901 20:23 and 0905 20:25). However, smaller surface values of CO around 200 ppbv are often overestimated at this airport during this period with large bias beyond a factor 2 (0901 15:25 and 0904 15:26). The maxima of CO observed at Malabo near 3000 and 4000 m are related to transport from forest fires occurring in Southern Africa during this period. CO values from these profiles are correlated to those of ozone (see ozone section). The corresponding maxima are detected by the models but often at a slightly lower altitude than observations. The magnitude of these maxima is well represented by the control run while the o-suite overestimates. In the UTLS, o-suite and control run behave similarly at all locations and agree well with observations.

North America

Three CO profiles are available at the airport of Miami during SON 2019 (Fig. 4.2.6). In these three profiles CO values of about 200 ppbv are observed near the surface. These values are largely underestimated by both, the o-suite and control run. Boundary layer values are also underestimated for two of these profiles (days: 1118 and 1130). In the UTLS the results of both runs are similar and models agree well with observations.
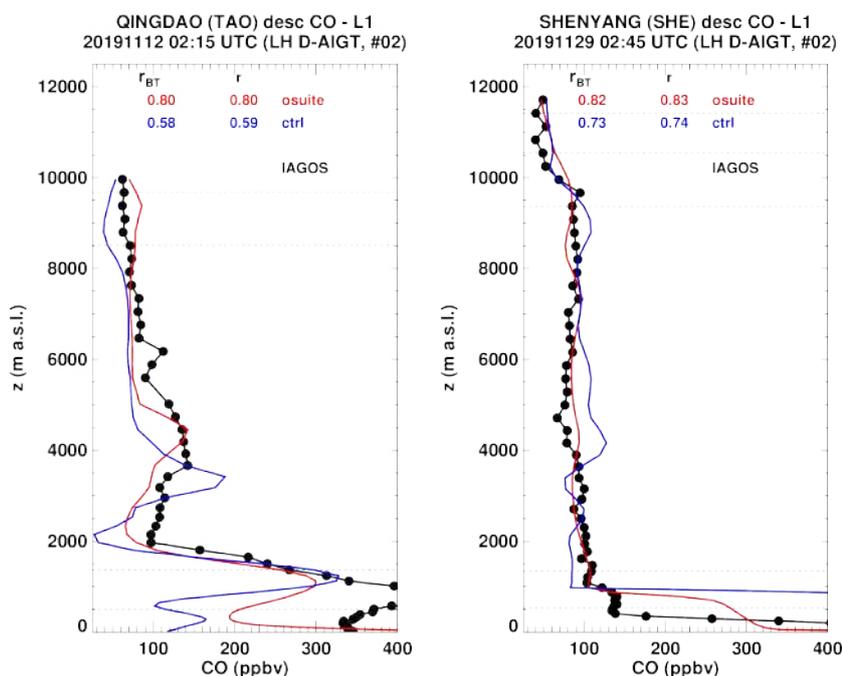
Figure 4.2.8: Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over East Asia during SON 2019.

Eastern Asia

Only two CO profiles are available at the Asian airports of: Qingdao and Shenyang (see Fig. 4.2.8). The high CO mixing ratios in the surface layer of both profiles and the complex shape of the profile at Qingdao are reproduced by both runs with a better performance of the o-suite as compared to the control run. In the lowest layers, the control run presents larger underestimations than the o-suite. In the free troposphere the results of the two runs are more similar than in the low troposphere with again a slightly better performance from the o-suite. In the UTLS the agreement is better than in the lowest layers and the performances of the models are similar.

Central America

During SON 2019, a few profiles of CO are available at Panama City and one in Cancun, with some examples in Fig. 3.2.9. At both airports CO values are largely underestimated in the surface and boundary layer (by a factor 2 at Panama near the surface) with slightly better results from the control run. In the free troposphere, CO values remain underestimated at Panama while the model and observations agree well at Cancun. In the UTLS, the models performs similarly well at both locations.
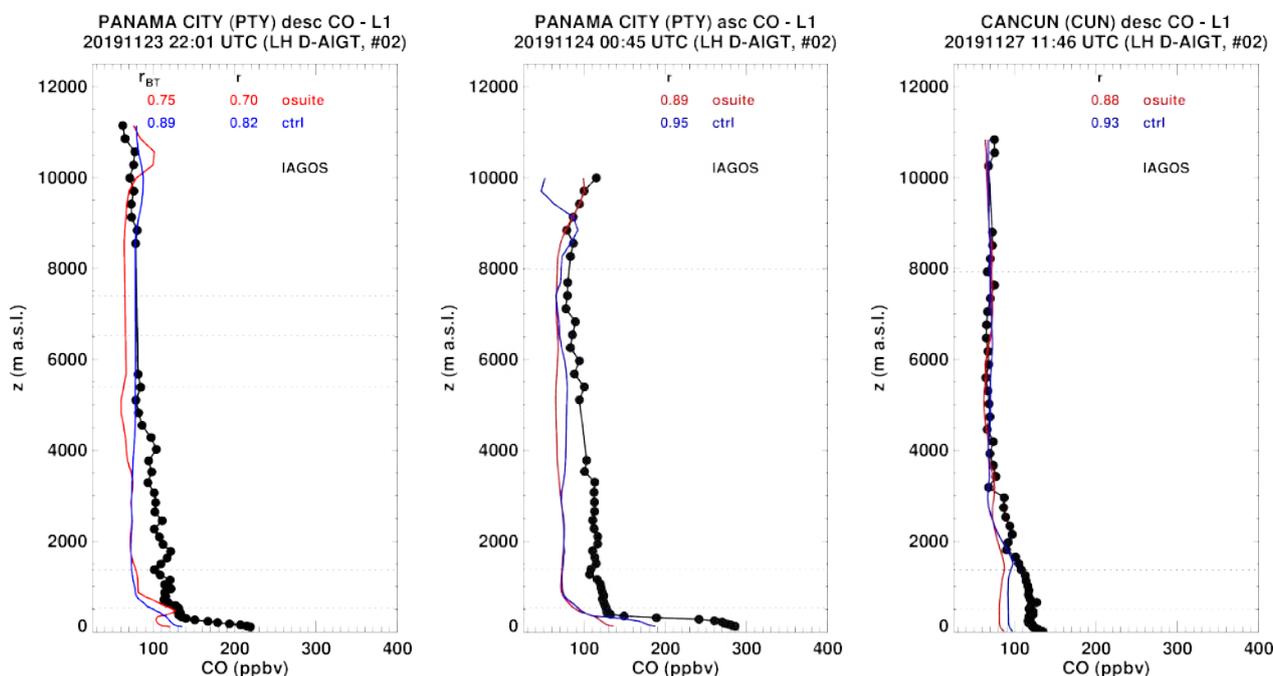
Figure 3.2.9 Daily profile for CO from IAGOS (black) and the two NRT runs (o-suite: red, control: blue) over South America during SON 2019.

4.3 Validation against FTIR observations from the NDACC network

In this section, we compare the CO profiles of the CAMS products with FTIR measurements at 21 FTIR stations within the NDACC network. These ground-based, remote-sensing instruments are sensitive to the CO abundance in the troposphere and lower stratosphere, i.e. between the surface and up to 20 km altitude. Tropospheric and stratospheric CO partial columns are validated. A description of the instruments and applied methodologies can be found at http://nors.aeronomie.be.

Figure 4.3.1 show that the o-suite tropospheric columns of CO agree well. The model upgrade (60 to 137 levels) implemented in July 2019 changes the overall biases in both the troposphere and stratosphere. The bias for the tropospheric columns becomes -7% in SON (-4% in JJA) and is larger than the reported measurement uncertainty. The stratospheric column bias reduces to +2% in SON (+6% in JJA) and now falls within the measurement's uncertainty.

Fig. 4.3.2 shows a trend in the tropospheric CO column at Jungfraujoch (4km – TP) of about 1.5% per year. A similar trend is observed at Zugspitze (3km asl), but not at other non-mountain sites like St Petersburg. The trend at the o-suite 1dFC at both mountain stations is much lower (around -0.5%/y), which suggests the trend is located in the upper tropospheric column and is related to the assimilation.

The Taylor diagrams in Figure 4.3.3 provide information on the correlation of all three CAMS products under consideration with the FTIR time series. Leaving out the sites with few measurements, the assimilation has a positive effect on the correlation coefficient. Looking at the correlation values for the period 2019 SON, the o-suite 1d FC (averaged correlation for all sites is 0.78) is comparable to the o-suite AN (averaged correlation for all sites is 0.79).
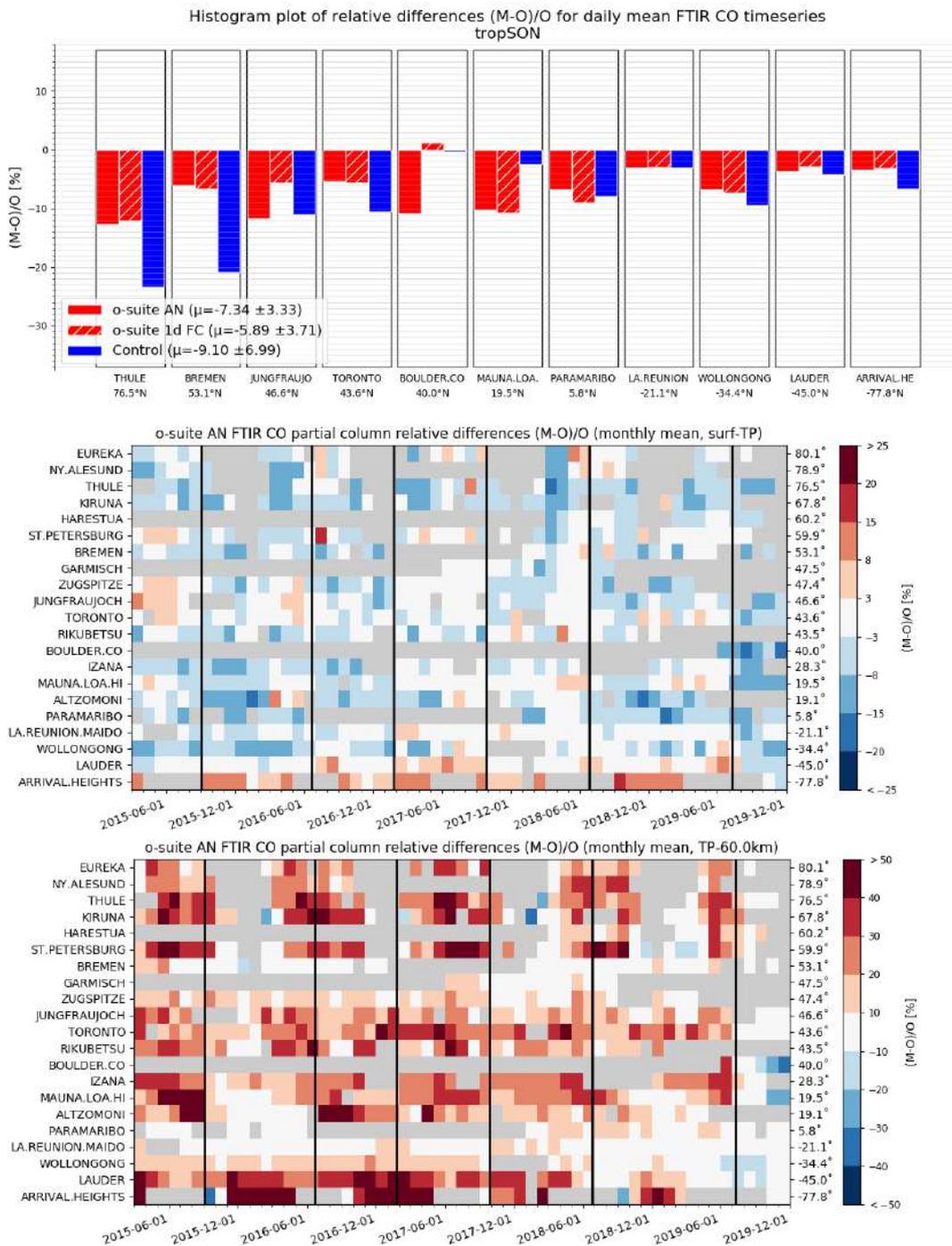
Figure 4.3.1: Seasonal relative mean bias for tropospheric CO columns (MB, %) for the considered period 2019 SON (top) and monthly mean biases for a longer time period for the tropospheric CO columns (middle) and stratospheric CO columns (bottom) (model upgrades are indicated in black vertical lines).The overall uncertainty for the CO measurements is approximately 3% on the tropospheric columns and 10% for the stratospheric columns. The o-suite analysis averaged bias in tropospheric columns increased to -7% for SON 2019. The bias in the stratosphere reduced to +2% and lies within the measurement's uncertainty. Stations are sorted with decreasing latitude (northern to southern hemisphere).
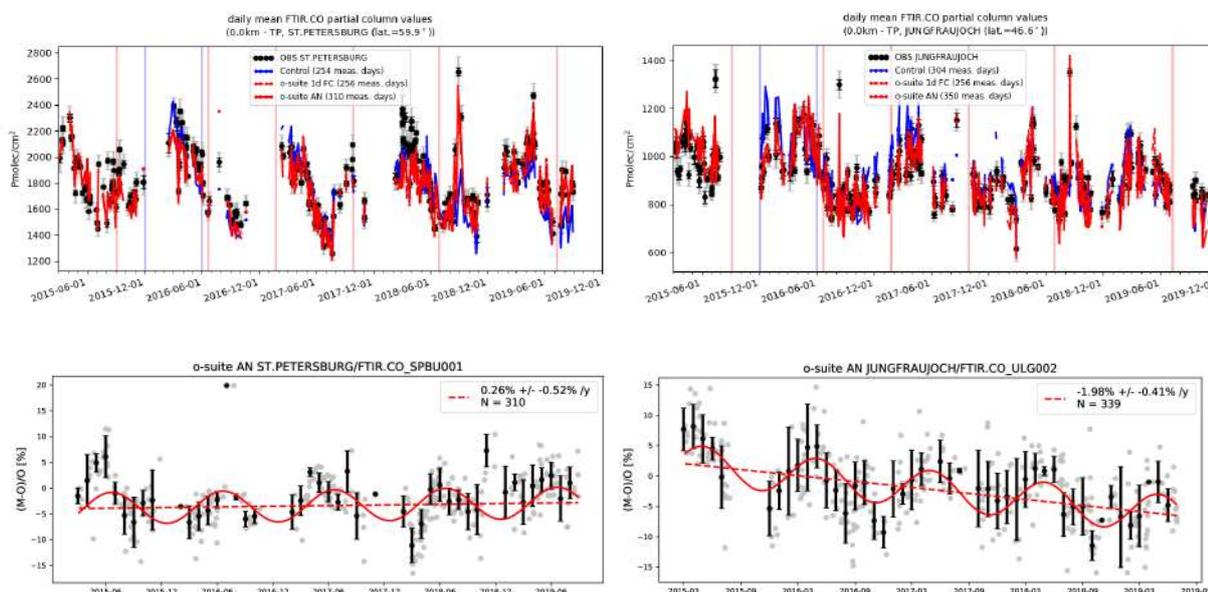
Figure 4.3.2: Top: daily mean values of tropospheric CO columns by the o-suite (AN and 1d FC, red) and the Control run (blue) compared to NDACC FTIR data at St Petersburg and Jungfraujoch for the period March 2015-August 2019. During March 2018 the o-suite underestimated the CO columns at St. Petersburg. Bottom row contains a linear fit and seasonal cycle fit through the relative differences for the o-suite analysis. An underestimation is observed during the local autumn/winter months. The negative trend at Jungfraujoch is not seen in the o-suite 1-day forecast.

4.4 Validation against FTIR observations from the TCCON network

CO column averaged mole fractions of the CAMS models are compared with data from the Total Carbon Column Observing Network (TCCON). Column averaged mole fractions provide different information content than the in-situ measurements and are therefore complementary to the in situ data. In this section, we compare column averaged mole fractions of CO of the CAMS models with TCCON retrievals. Data from the following TCCON sites has been used:

Izana (Blumenstock et al., 2017), Reunion (De Mazière et al., 2017), Bialystok (Deutscher et al., 2017), Manaus (Dubey et al., 2017), Four Corners (Dubey et al., 2017), Ascension (Feist et al., 2017), Anmeyondo (Goo et al., 2017), Darwin (Griffith et al., 2017), Wollongong (Griffith et al., 2017), Karlsruhe (Hase et al., 2017), Edwards (Iraci et al., 2017), Indianapolis (Iraci et al., 2017), Saga (Kawakami et al., 2017), Sodankyla (Kivi et al., 2017), Hefei (Liu et al., 2018), Tsukuba (Morino et al., 2017), Burgos (Morino et al., 2018), Rikubetsu (Morino et al., 2017), Bremen (Notholt et al., 2017), Spitsbergen (Notholt et al., 2017), Lauder (Sherlock et al., 2017, Pollard et al., 2019), Eureka (Strong et al., 2018), Garmisch (Sussmann et al., 2017), Zugspitze (Sussmann et al., 2018), Paris (Te et al., 2017), Orleans (Warneke et al., 2017), Park Falls (Wennberg et al., 2017), Caltech (Wennberg et al., 2017), Lamont (Wennberg et al., 2017), Jet Propulsion Laboratory (Wennberg et al., 2017), East Trout Lake (Wunch et al., 2017)

Figure 4.3.3: Taylor diagrams relating the standard deviations for the model /GB time series of tropospheric CO column data and their correlation. All timeseries are normalized such that the std of the model is 1. The variability of the CO columns in the o-suite 1dFC deviates more from the variability in the FTIR columns compared to the o-suite AN. At the end of November, the Wollongong measurements already show increased values due to start of the forest fires in Australia and explains why the std of the measurements is higher than the model.

For the validation of the models in September, October and November data is available from 12 TCCON sites. The sites that provide data for the full time period (Sep-Nov) are Orleans, Nicosia (Cyprus) and Reunion. The site Nicosia (Cyprus) is a new site, operational since September 2019. It is operated in collaboration between the Cyprus Institute and the University of Bremen.

The comparisons show that the last model update improved the comparison between the o-suite AN and o-suite 1d FC models with TCCON data, while the control model shows a low bias (Fig. 4.4.1 and 4.4.2). Also, the variation of the CO after the model update is captured well by the o-suite AN and o-suite 1d FC models, as shown exemplary for the TCCON site Orleans (Fig 4.4.3). The available data for the reporting period is limited to the Northern Hemisphere, except for the site Reunion.
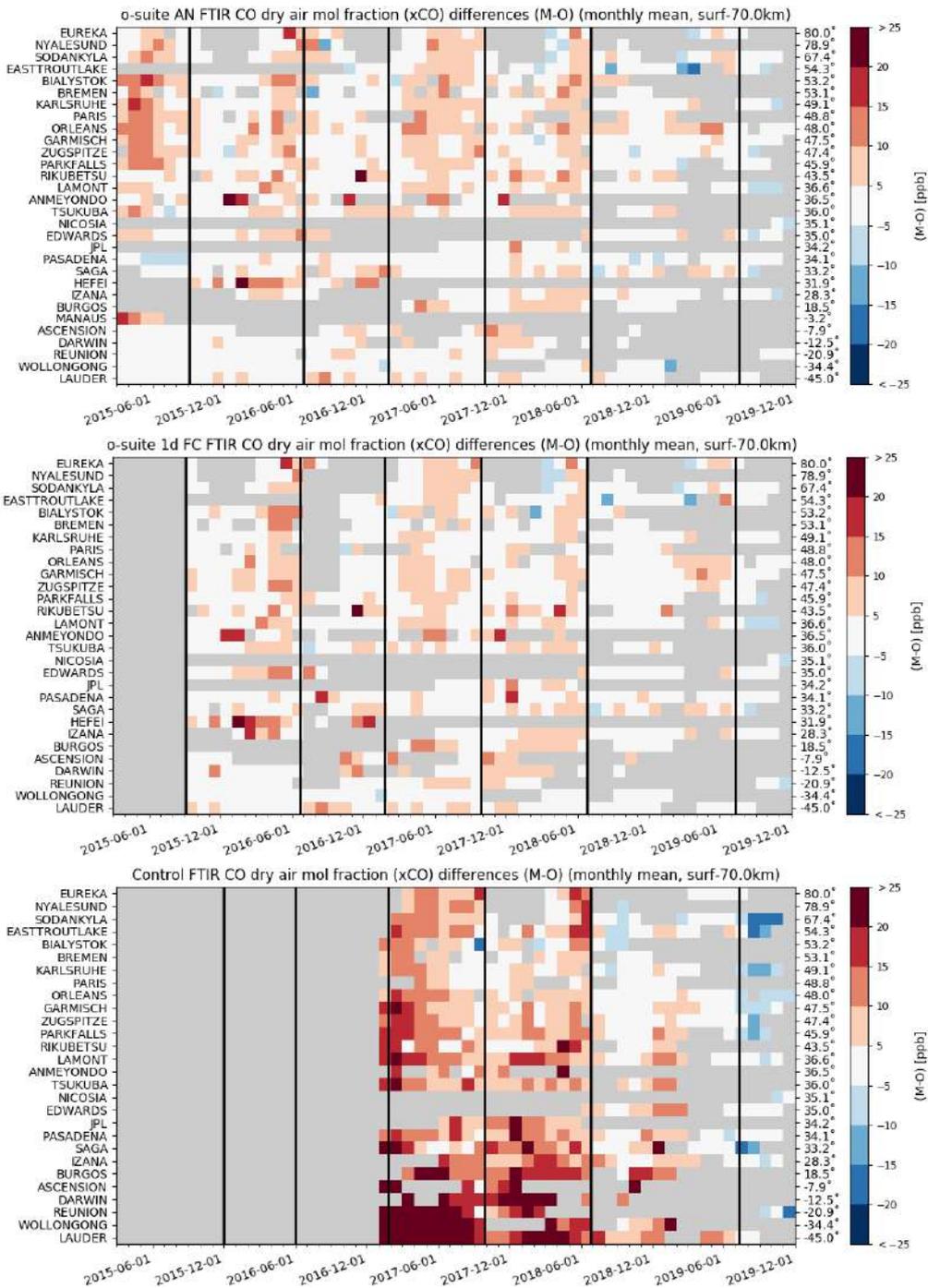
Figure 4.4.1: Monthly differences for the last 4.5 years. The stations are sorted by latitude (northern to southern hemisphere). The upper plot shows the comparison with the o-suite AN, the middle the comparison with the o-suite 1d FC and the lower plot the comparison with the control model.
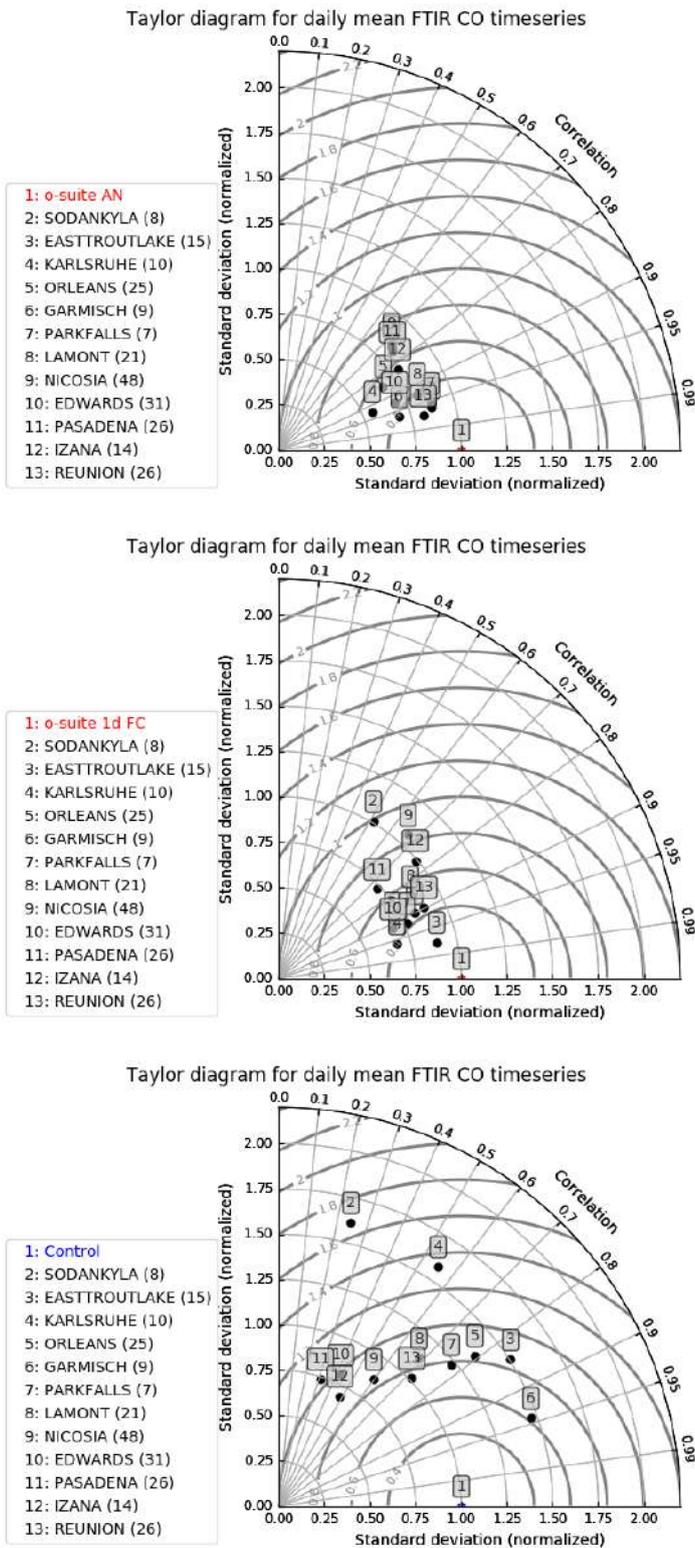
Figure 4.4.2: Taylor diagrams for the reporting period Sep-Nov 2019. The upper plot shows the comparison with the o-suite AN, the middle the comparison with the o-suite 1d FC and the lower plot the comparison with the control model.
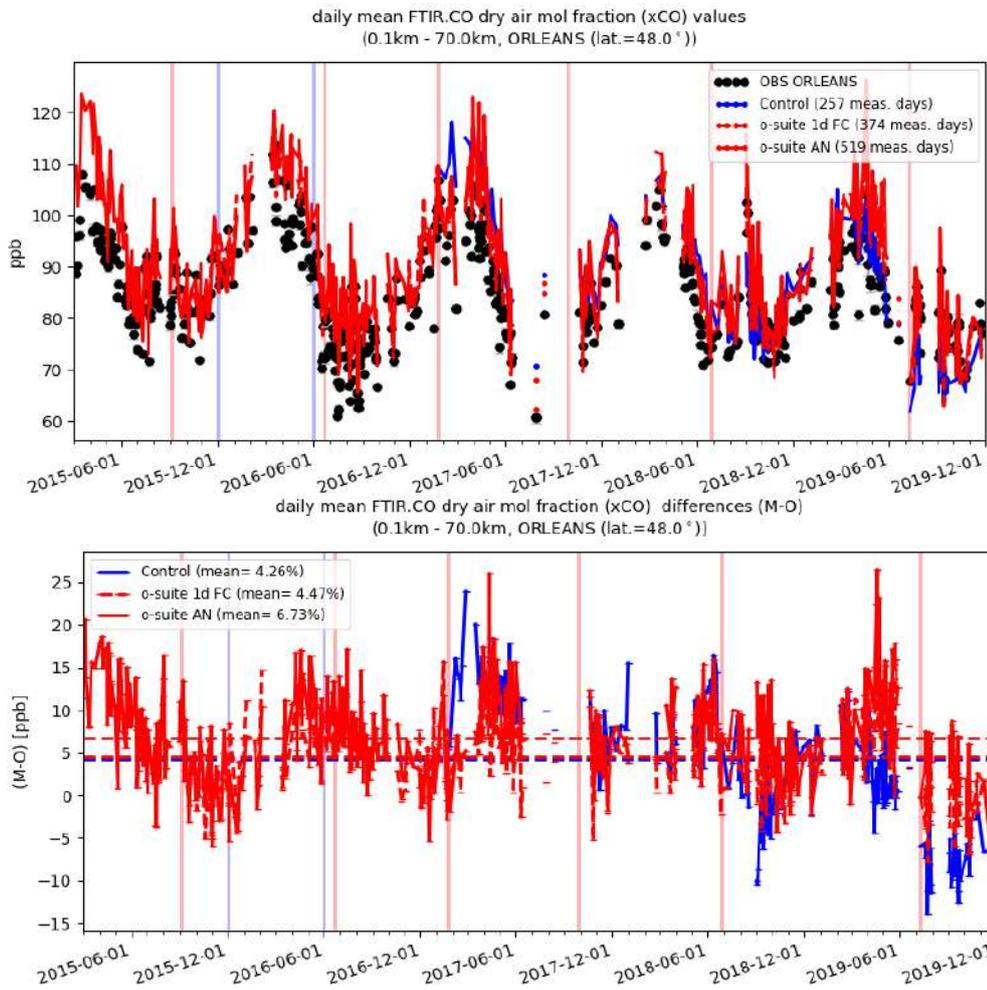
Figure 4.4.3: Comparison of the CO model data with TCCON CO at Orleans.
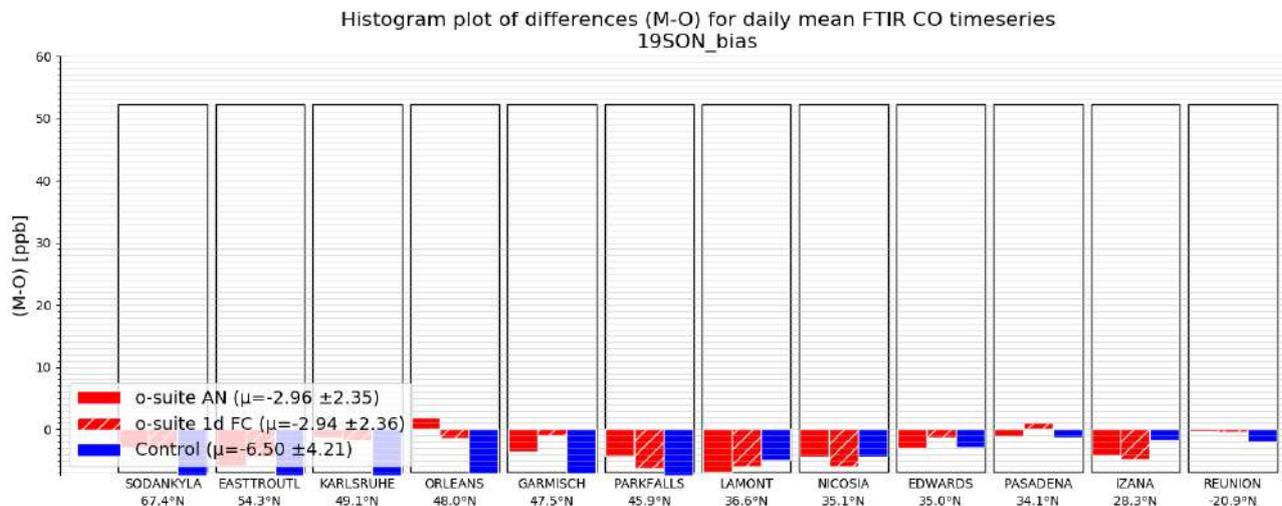
Figure 4.4.4: Differences during the reporting period. The different sites cover different periods of the comparison period. Therefore, only a qualitative comparison can be done.

4.5 Evaluation with MOPITT and IASI data

In this section, modelled CO total columns are compared to MOPITT version 8 (thermal infrared radiances) (Emmons et. al., 2009, Deeter et al., 2010) and IASI satellite retrievals (Clerbaux et al., 2009). Figure 4.5.1 shows the global distribution of CO total columns retrieved from MOPITT V8 (top left) and IASI (top right) and the relative biases of the model runs with respect to MOPITT V8 for September 2019.

MOPITT shows high values over the biomass burning area in Central Africa, South America and over the north-eastern part of China and Indonesia. IASI shows higher values over the abovementioned regions and also high values over the Bay of Bengal south of China.

The modelled CO geographical distribution and magnitude of values show that the model performs reasonably well. The relative difference between the model runs and MOPITT shows that the o-suite performs better than the control run without data assimilation. The o-suite is generally underestimating the satellite data by about 10% with some regional exceptions where the negative bias reaches 20% (mostly over the land). The control run shows an underestimation of the satellite data over the Northern Hemisphere continents and south of Africa up to 30% and over the Southern Hemisphere by up to 20% and overestimation over the tropics by up to 40%. The o-suite run shows a growing positive bias on the 4th forecast day over the fire active areas in Indonesia and Central Africa, as well as a growing negative bias over the continents, especially over the south of Africa.

Figure 4.5.2 shows time series of CO total column data for MOPITT V8, IASI and the model runs over the eight selected regions. For the comparison with MOPITT, the modelled CO concentrations were transformed using MOPITT V8 averaging kernels (Deeter, 2004). Both, MOPITT and IASI CO total columns are assimilated in the o-suite run, while a bias correction scheme is applied to IASI data to bring it in line with MOPITT. MOPITT and IASI CO total columns show a relatively similar variability over different regions. IASI CO values are lower than MOPITT over most regions with some seasonal exceptions until the year 2016. Since then IASI and MOPITT are more consistent with each other over Europe, the US and East Asia. Significant difference between MOPITT and IASI are observed over the Alaskan and Siberian fire regions in winter seasons, with IASI CO total column values being lower up to 30%. In North and South Africa, deviations become larger since 2016 with IASI values being higher than MOPITT by up to 20%. The modelled seasonality of CO total columns is in relatively good agreement with the retrievals. In general, the comparison between the o-suite and control run shows that the assimilation of satellite CO has a more positive, pronounced impact on model results over East and South Asia, South Africa, and since the end of 2016, over the US in winter and spring seasons, and smaller impact over the other regions. Since June 2016, the o-suite shows very good agreement with the satellite retrievals over Europe and the US with biases less than 5%. In late summer and early autumn of 2018 over Europe, the control run has larger negative biases compared to the satellite data then early in 2018 and the two previous autumn seasons.

A general reduction of CO values from the year 2015 to the year 2018 can be seen over Europe, the US and East Asian regions. The South African region shows a slight increase of the seasonal minimum compared to previous springs.
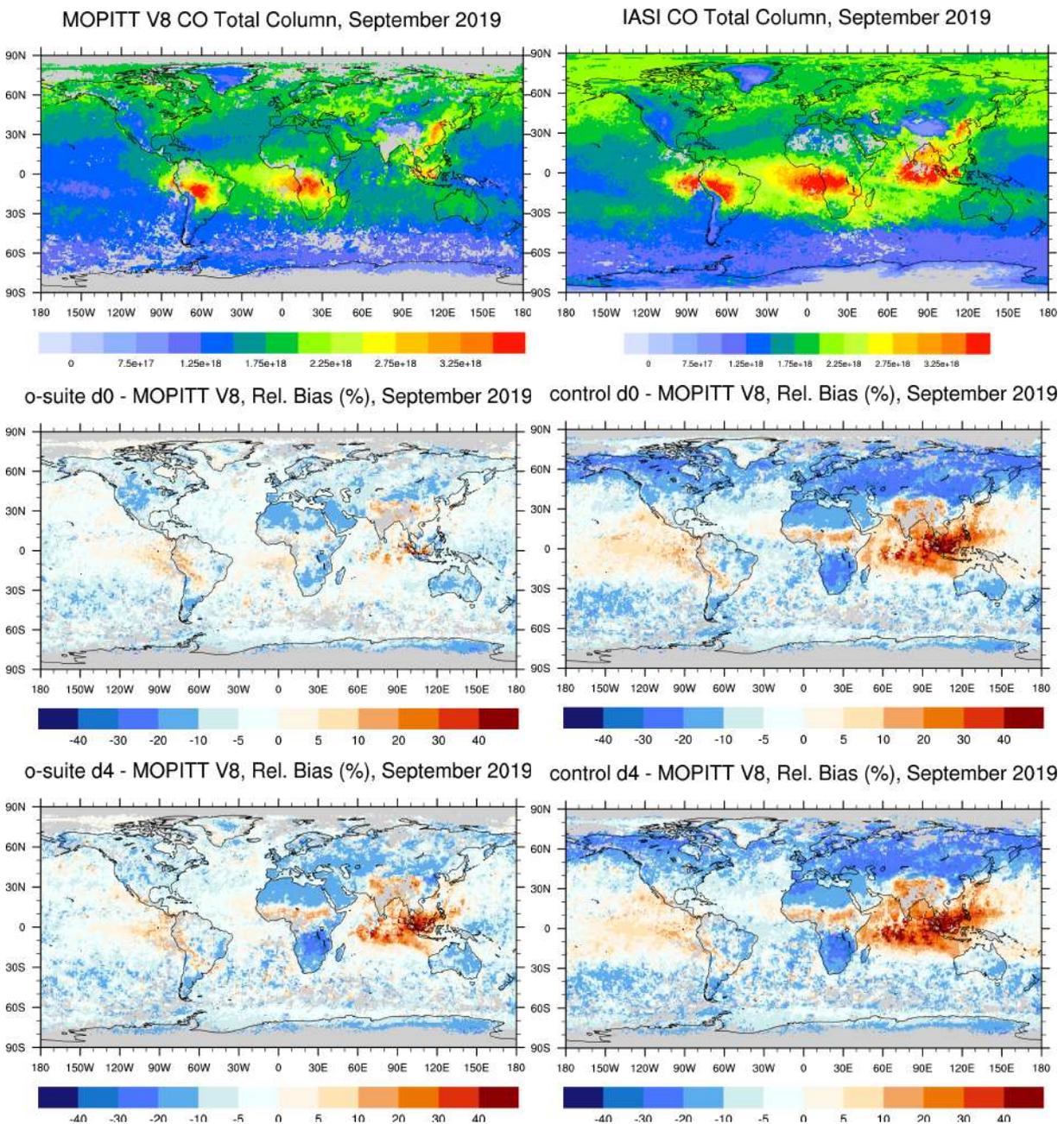
Fig. 4.5.1: CO total columns for MOPITT V8 (top left) and IASI (top right) satellite retrievals and relative difference between the model runs and MOPITT for September 2019: o-suite (middle left), control run (middle right), o-suite 4th forecast day (bottom left), o-suite 4th forecast day (bottom right). Grey colour indicates missing values.

Summer 2019 was characterised by strong fire events in Siberia. This can be seen in IASI data (peak in August), but it is not reflected in the MOPITT data partly due to only few days of observations available in August.

The modified normalized mean bias (MNMB) of the model runs compared to MOPITT V8 (Fig. 4.5.3) allows quantifying the impact of the assimilation on the model performance. The o-suite model run shows negative biases over Europe, the US and Alaskan fire regions with some seasonal exceptions.
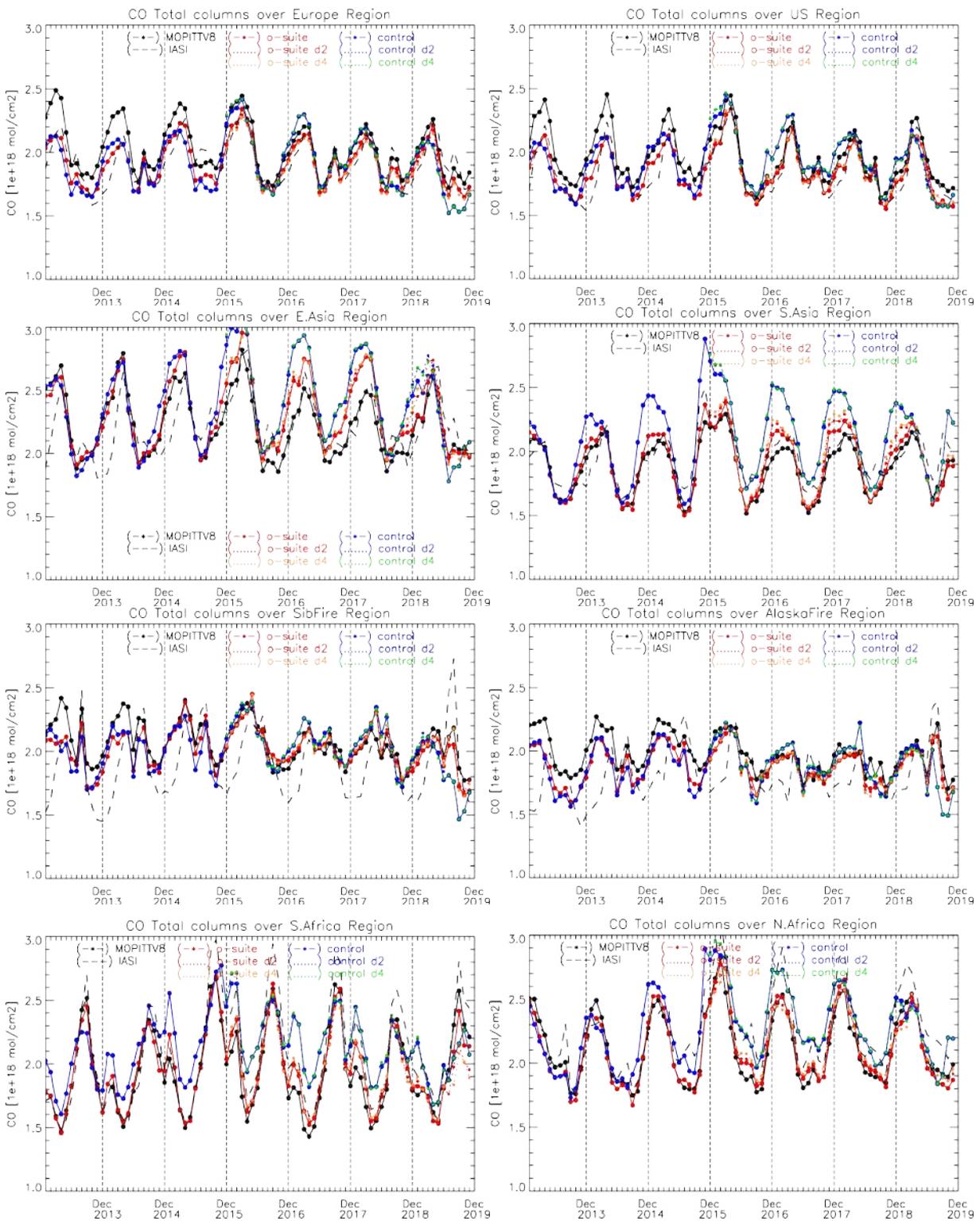
Fig. 4.5.2: Time series of CO total columns for satellite retrievals MOPITT V8, IASI (black) and the model runs over the selected regions: o-suite (red, solid), control (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted), control 2nd forecast day (blue, dotted), control 4th forecast day (green, dotted). Period: January 2013 to November 2019.
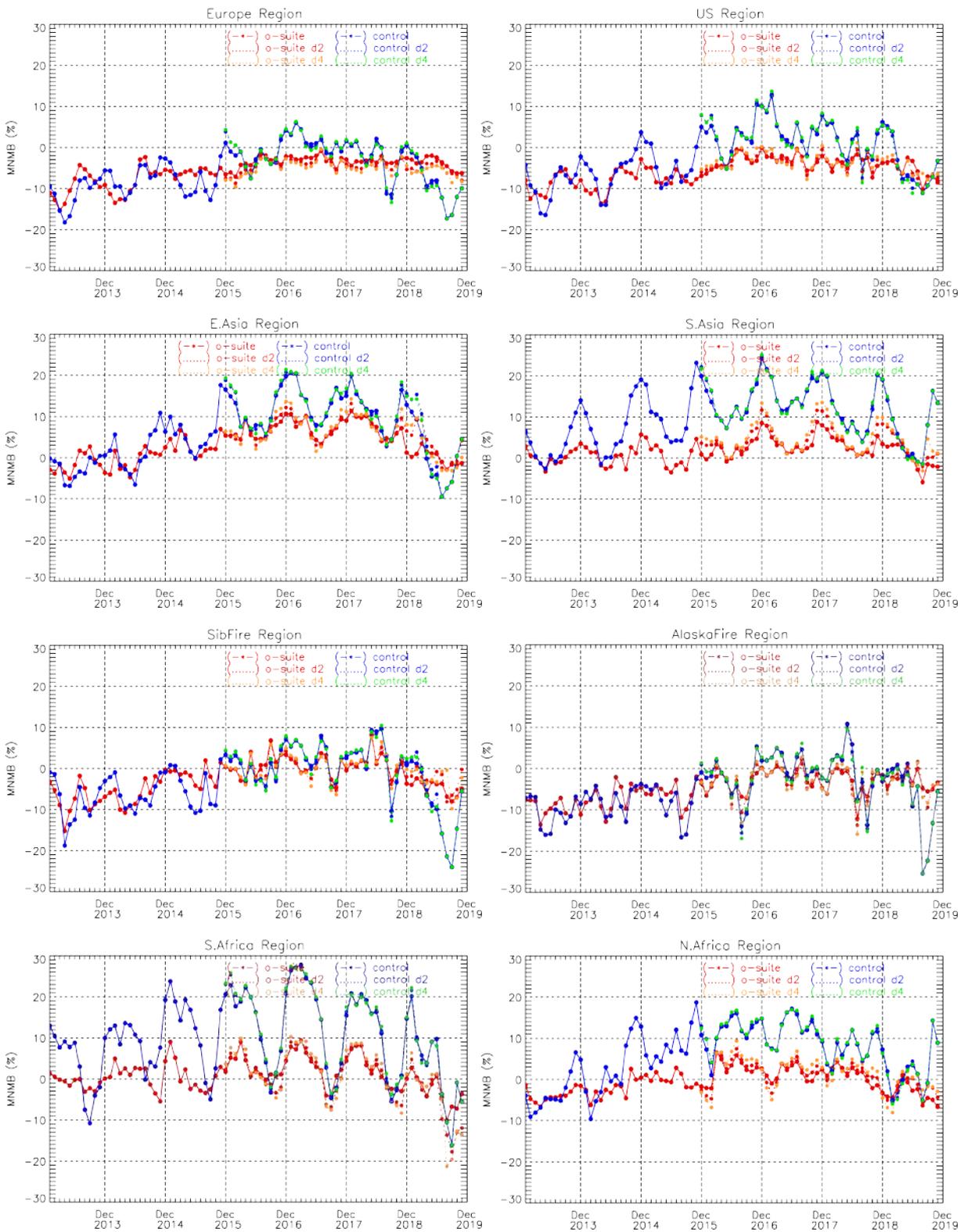
Fig. 4.5.3: Timeseries of modified normalized mean bias (%) for CO total columns from the model simulations vs MOPITT V8 retrievals over selected regions. O-suite (red, solid), control run (blue, solid), o-suite 2nd forecast day (red, dotted), o-suite 4th forecast day (orange, dotted), control 2nd forecast day (blue, dotted), control 4th forecast day (green, dotted). Period: January 2013 to November 2019.

The control run shows a systematic positive bias up to 20% over South Asia in November-December 2014, 2015, 2016, and 2017. Over southern Africa, the control run overestimates satellite retrieved values by up to 25% in winter and spring 2015, 2016, and 2017. In general, the o-suite is within +/- 10% in all regions, while the control run shows larger biases over East and South Asia and North and South Africa, as well as stronger seasonal cycles.

In the second half of the year 2019, we see the following changes:

- A slight increase of negative biases over Europe in the o-suite run and more pronounced underestimations in the control run.

- Enhanced negative biases over the US for both CAMS configurations.

- Improvement of the o-suite results over East Asia (bias is almost zero) and strong reduction of bias in the control run.

- Reduction of biases over South Asia for both runs.

- General change of bias sign in the control run from positive to negative over the Siberian fire region and enhanced negative biases in the o-suite run.

- Strong increase of negative bias in the control run over the Alaskan region.

- Stronger underestimation over South Africa in both, the o-suite and control run.

- Stronger underestimation in the o-suite run over North Africa.

- In general, the increased underestimation in the o-suite run can be seen over most of the regions.

5. Tropospheric nitrogen dioxide

5.1 Evaluation against GOME-2 and TROPOMI retrievals

In this section, model columns of tropospheric NO_2 are compared to SCIAMACHY/Envisat NO_2 satellite retrievals (IUP-UB v0.7) [Richter et al., 2005] for model data before April 2012, and to GOME-2/MetOp-A NO_2 satellite retrievals (IUP-UB v1.0) [Richter et al., 2011] for more recent simulations. First comparisons to TROPOMI/Sentinel-5P data (IUP-UB v0.1, preliminary) are provided, using the CAMS o-suite as a-priori in the TROPOMI retrievals. This satellite data provides excellent coverage in space and time and very good statistics. However, only integrated tropospheric columns are available, and the satellite data is always taken at the same local time, roughly 09:30 LT for GOME-2, 10:00 LT for SCIAMACHY and 13:30 LT for TROPOMI and at clear sky only. Therefore, model data are vertically integrated, interpolated in time and then sampled to match the satellite data. The satellite data were gridded to model resolution (currently 0.4° x 0.4° degree). Model data were treated with the same reference sector subtraction approach as the satellite data for all SCIAMACHY/GOME-2 comparisons. For all comparisons to TROPOMI satellite data, tropospheric NO_2 columns over the clean Pacific reference sector simulated by CAMS-global were added to the TROPOMI data, so that the comparison of absolute columns is accomplished. For TROPOMI comparisons before July 2019, the stratospheric contribution has been removed from the measurements according to the method by Hilboll et al. (2013) using simulations from the B3D-CTM (Sinnhuber at al., 2003a; Sinnhuber et al., 2003b; Winkler et al., 2008) scaled to satellite values over the clean Pacific reference sector. Since July 2019, the reference sector method has been applied to the TROPOMI data. Uncertainties in NO_2 satellite retrievals are large and depend on the region and season. Winter values in mid and high latitudes are usually associated with larger error margins. Systematic uncertainties in regions with significant pollution are on the order of 20% – 30%.

Figure 5.1.1 shows global maps of GOME-2 and model monthly mean tropospheric NO_2 columns as well as differences between retrievals and simulations for Sep 2019 as an example of the maps for autumn 2019. The overall spatial distribution and magnitude of tropospheric NO_2 is well reproduced by both CAMS runs, indicating that emission patterns and NO_x photochemistry are reasonably well represented. Some differences are apparent between observations and simulations, with generally larger shipping signals simulated by the models. For example, shipping signals are much more pronounced in model simulations to the south of India. Emissions over Europe and especially the pollution hotspots around the Benelux countries are regularly underestimated, especially during winter. However, other local maxima of tropospheric NO_2 observed over anthropogenic emission hotspots in East Asia (e.g. over the heavily populated Sichuan Basin; 30°N, 105°E), India and others such as Teheran, Mecca, around Lebanon/Israel and Moscow and over boreal forest fires (see Alaska Fig. 5.1.1) are regularly overestimated. Likewise, values over the Persian Gulf and the Red Sea are regularly overestimated (mainly summer and autumn). Similar conclusions arise from TROPOMI based map comparisons shown in Figure 5.1.2, though the magnitude of the differences between observations and simulations differs. The large differences at high southern latitudes are due to the imperfections of the reference sector method used for subtraction of the stratosphere and should be ignored. Differences in comparison results are in principal due to differences in observation time or differences in the retrieval products.
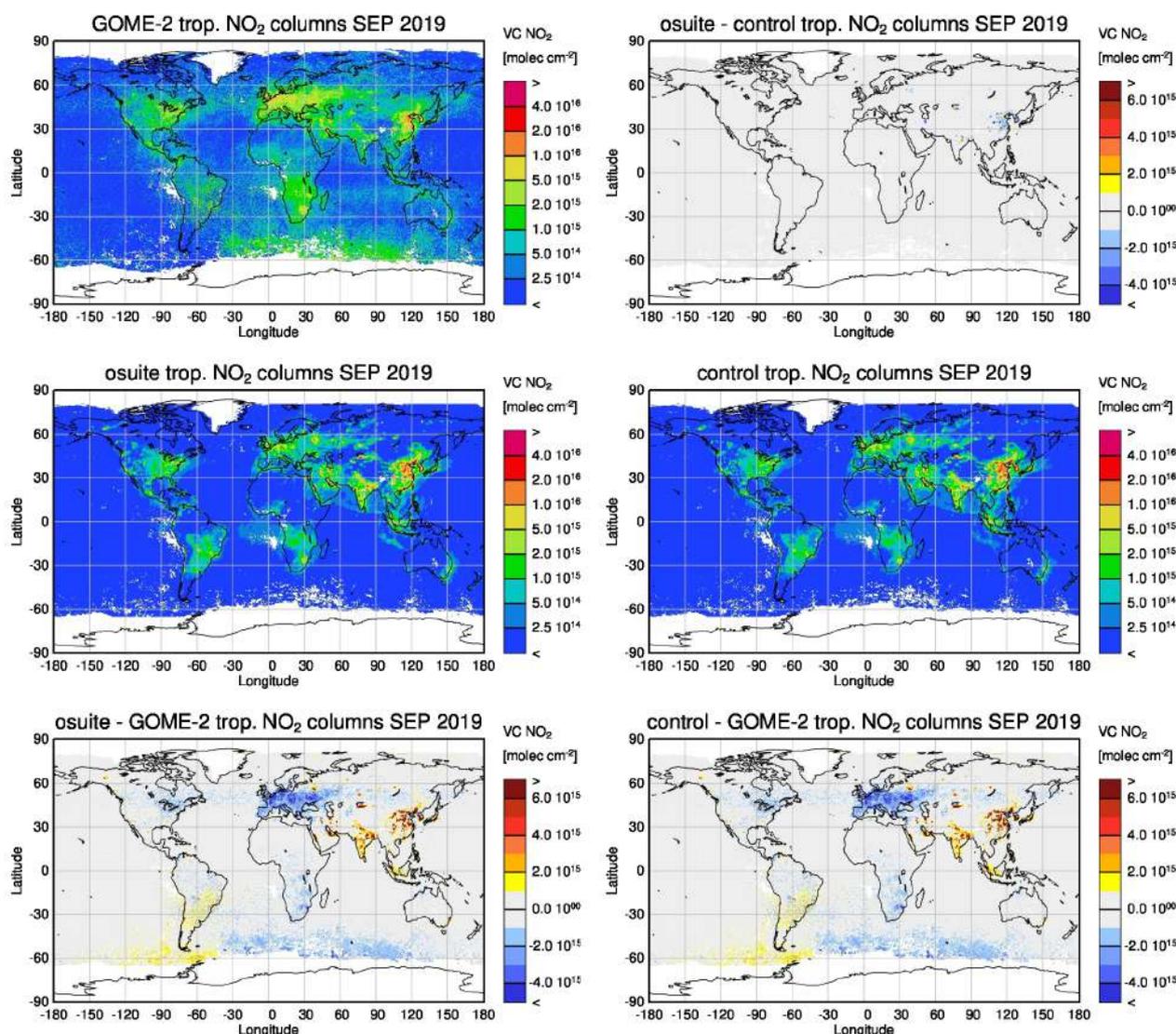
Figure 5.1.1: Global map comparisons of satellite retrieved, and model simulated tropospheric NO$_2$ columns [molecules cm^{-2}] for Sep 2019. The top row shows monthly mean tropospheric NO$_2$ columns retrieved from GOME-2 as well as the difference between o-suite and control, the second row shows the corresponding tropospheric NO$_2$ columns for model simulated averages. The third row shows differences of monthly means between models and GOME-2. GOME-2 data were gridded to model resolution (i.e. 0.4° x 0.4° degree). Model data were treated using the same stratospheric correction method as for the satellite data.

Closer inspection of the seasonal variation of tropospheric NO$_2$ in some selected regions (Fig. 5.1.3) reveals significant differences between measurements and model results and points to some simulation problems. Over regions where anthropogenic emissions are major contributors to NO$_x$ emissions, models catch the shape of the satellite time series rather well. However, over East-Asia absolute values and seasonality were strongly underestimated before 2014 by all model runs (most likely due to an underestimation of anthropogenic emissions) for all seasons apart from summertime minima, with the o-suite showing the best results since an upgrade in July 2012. As wintertime NO$_2$ column retrievals decreased significantly in 2014, model simulated wintertime maxima are in better agreement with the satellite retrieved ones for recent years. However, the
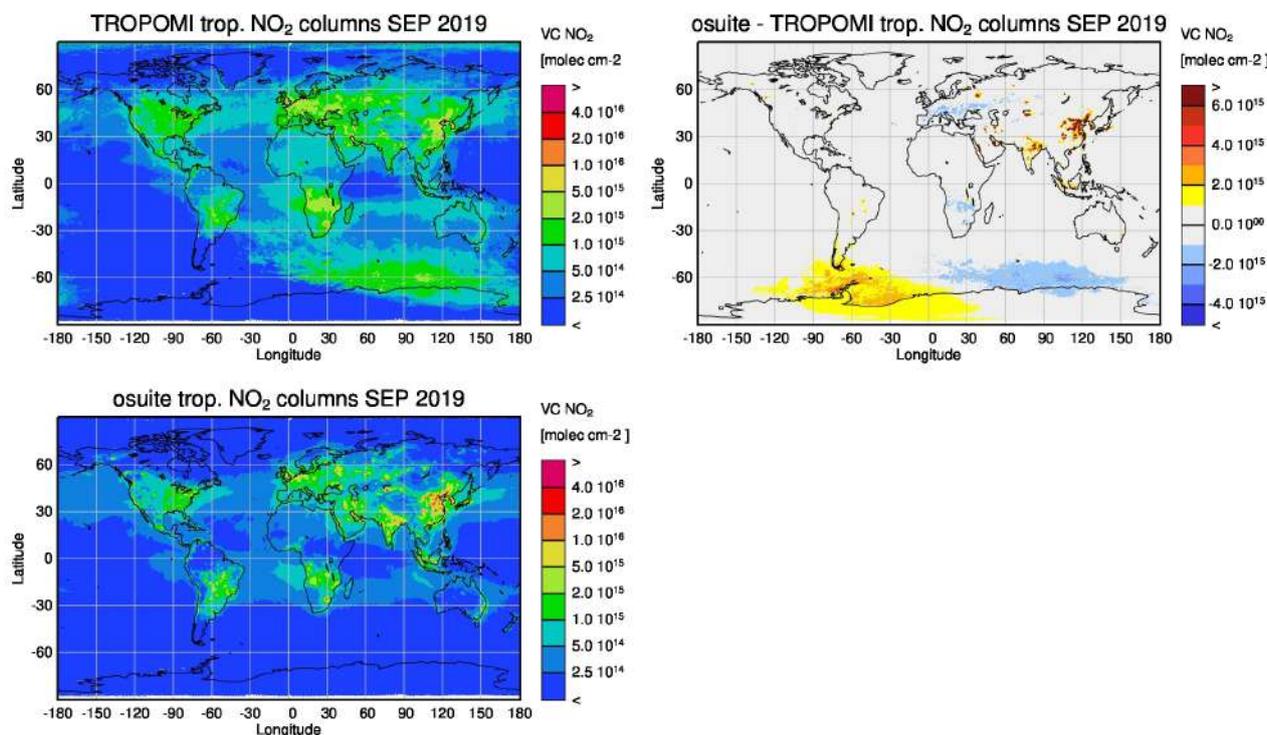
Figure 5.1.2: Global map comparisons of satellite retrieved, and model simulated tropospheric NO$_2$ columns [molecules cm-2] for Sep 2019: (top left) TROPOMI, (top right) o-suite minus TROPOMI, (bottom left) o-suite. TROPOMI data were gridded to model resolution (i.e. 0.4° x 0.4° degree) and the CAMS o-suite was used as a-priori in the retrievals. Comparisons to the control are not available for this report.

observed NO$_2$ decrease is not reproduced by the simulations and therefore the better agreement for more recent years cannot be attributed to model improvements. Moreover, summertime model minima increased in 2015 compared to previous years, which is in contrast to the satellite retrievals, so that the simulated values for the summers since 2015 are about 50% larger than satellite retrieved ones.

As for East-Asia, a decrease in satellite retrieved values also occurred in 2015 over Europe where a peak is usually found around January, which was, as a result, only slightly underestimated by the models for January 2015. The underestimation of tropospheric NO$_2$ columns over Europe may be caused to some extent by a change of emission inventories in 2012. However, the situation changed for the three winter periods between 2015 and 2017, for which GOME-2 shows (compared to previous years) a strong increase in January peak values, combined with a decrease in values for December and February that is not reproduced by the models. It is not clear if the GOME-2 observations are realistic here, although an inspection of daily GOME-2 satellite images did not point to problems regarding the retrieval. The retrievals show the same pattern as the simulations however for winter 2018/2019.

Over regions where biomass burning is the major contributor to NO$_x$ emissions, seasonality and amplitude of model columns are determined by fire emissions. The seasonality for the two regions in Africa was simulated reasonably well for 2010 and after October 2011. In the time period in between, a bug in reading fire emissions lead to simulation errors for all MOZART runs. Over North-
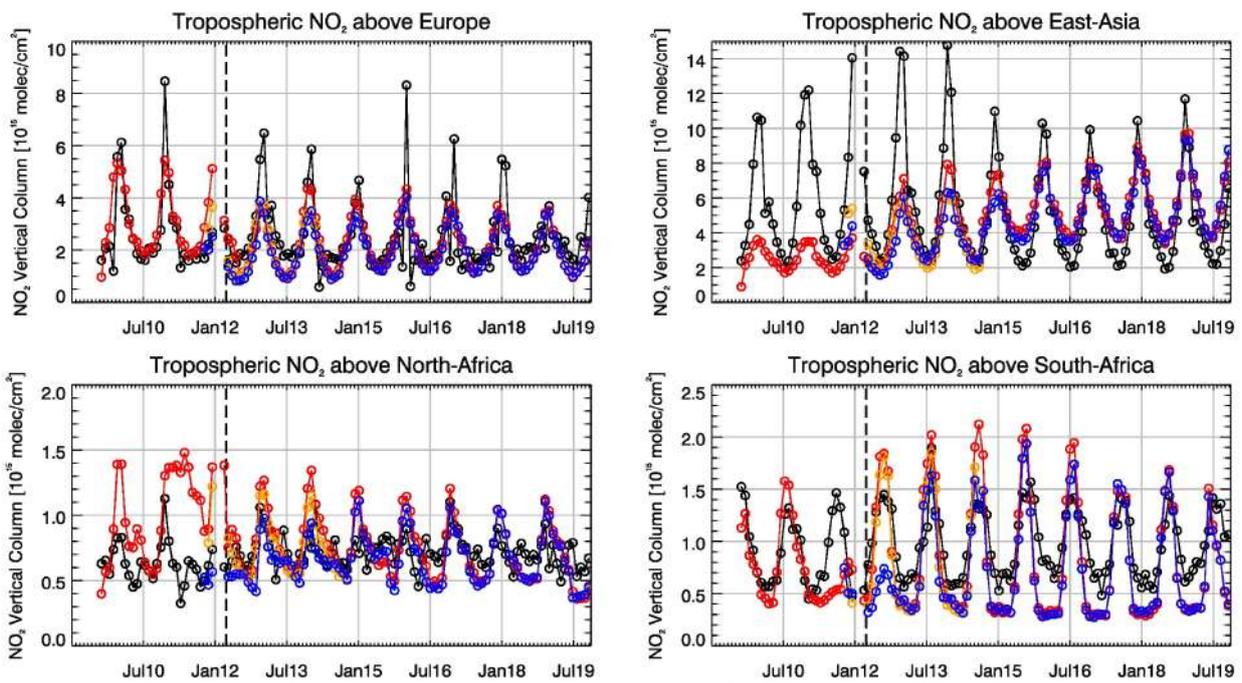
Figure 5.1.3: Time series of average tropospheric NO$_2$ columns [10^{15} molec cm^{-2}] from SCIAMACHY (up to March 2012, black) and GOME-2 (from April 2012 onwards, black) compared to model results (red: o-suite, blue: MACC_fncrt_TM5/MACC_CIFS_TM5/control, orange - MACC_fcnrt_MOZ) for different regions (see Annex 2 for definition of regions). The upper panels represent regions dominated by anthropogenic emissions, and the lower panels represent those dominated by biomass burning. Vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.
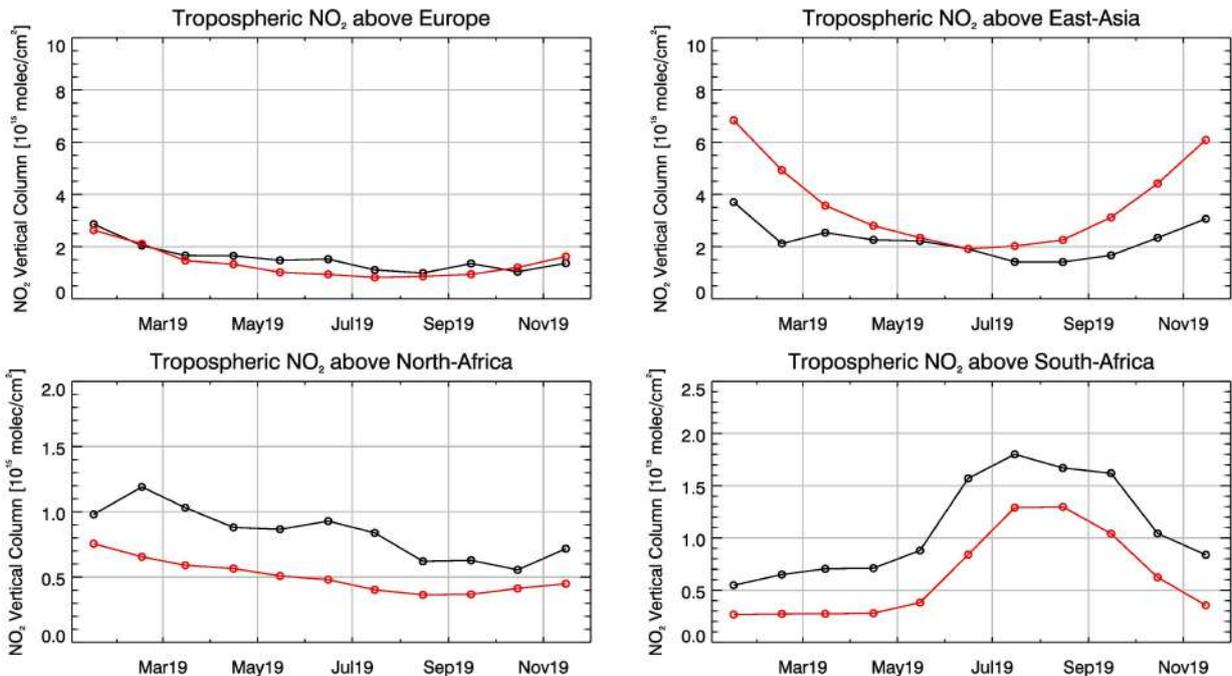
Figure 5.1.4: Time series of average tropospheric NO$_2$ columns [10^{15} molec cm^{-2}] from (black) TROPOMI compared to (red) o-suite model results since Jan 2019 (see Annex 2 for definition of regions). The upper panels represent regions dominated by anthropogenic emissions, and the lower panels represent those dominated by biomass burning.

Africa, the o-suite shows improved results since an update in July 2012 and the change to IFS-CB05 in September 2014. However, tropospheric NO_2 columns around December are still overestimated by the models. Summer to autumn NO_2 columns over North-Africa are underestimated compared to the satellite data from 2015 onwards and especially for 2019. The models (especially the o-suite) generally overestimate the seasonal cycle for South-Africa, particularly for 2014-2016 with an overestimation of the seasonal maximum, which usually occurs around August (e.g. by a factor of 1.4 larger compared to GOME-2 retrievals in 2016). However, August maxima are in better agreement since the upgrade of the o-suite in 2017, but minima during SH summer remain underestimated.

Time series comparisons between the o-suite and TROPOMI are shown in Figure 5.1.4 since January 2019. They show some differences with respect to the GOME-2 based ones: the o-suite overestimates values over East-Asia for Jan 2019, underestimates values over the African regions during the whole period of Jan-Aug 2019 (the seasonal cycle over South-Africa is not overestimated in this case) and simulations are close to the satellite observations over Europe according to the TROPOMI based comparisons. Differences in comparison results are in principal due to differences in observation time or differences in the retrieval products.

More NO_2 evaluation plots can be found on the CAMS website, see table 1.2.

5.2 Evaluation against ground-based DOAS observations

In this section, we compare the NO_2 columns of the CAMS products with UVVIS DOAS profile measurements at Uccle and column data from the other stations.[1] This ground-based, remote-sensing instrument is sensitive to the NO_2 abundance in the lower troposphere, up to 1km altitude with an estimated uncertainty of 8%. Tropospheric NO_2 profiles and columns are validated (up to 3.5km or 10km). A description of the instruments and applied methodologies is the same for all DOAS OFFAXIS measurements, see http://nors.aeronomie.be. It is important to mention here that the model partial column values are calculated from the smoothed model profiles. This guarantees that the model levels where the measurement is not sensitive do not contribute to the observed bias. We should mention that the measurement data is still catalogued as rapid delivery and not in the consolidated NDACC database.

Fig. 5.2.1 shows the biases for the latest validation periods June-August 2019 and Sep-Nov 2019 at the different sites. The bias has slightly decreased at De Bilt and Cabauw and Uccle but did not change significantly between both reporting periods. At the urban sites at Uccle and Athens a strong underestimation is observed. For the other sites (Bremen, De Bilt and Cabauw) the o-suite AN is able to capture only few of the high pollution events.

[1] No contribution from Xianghe, Reunion and OHP due to instrument failure.
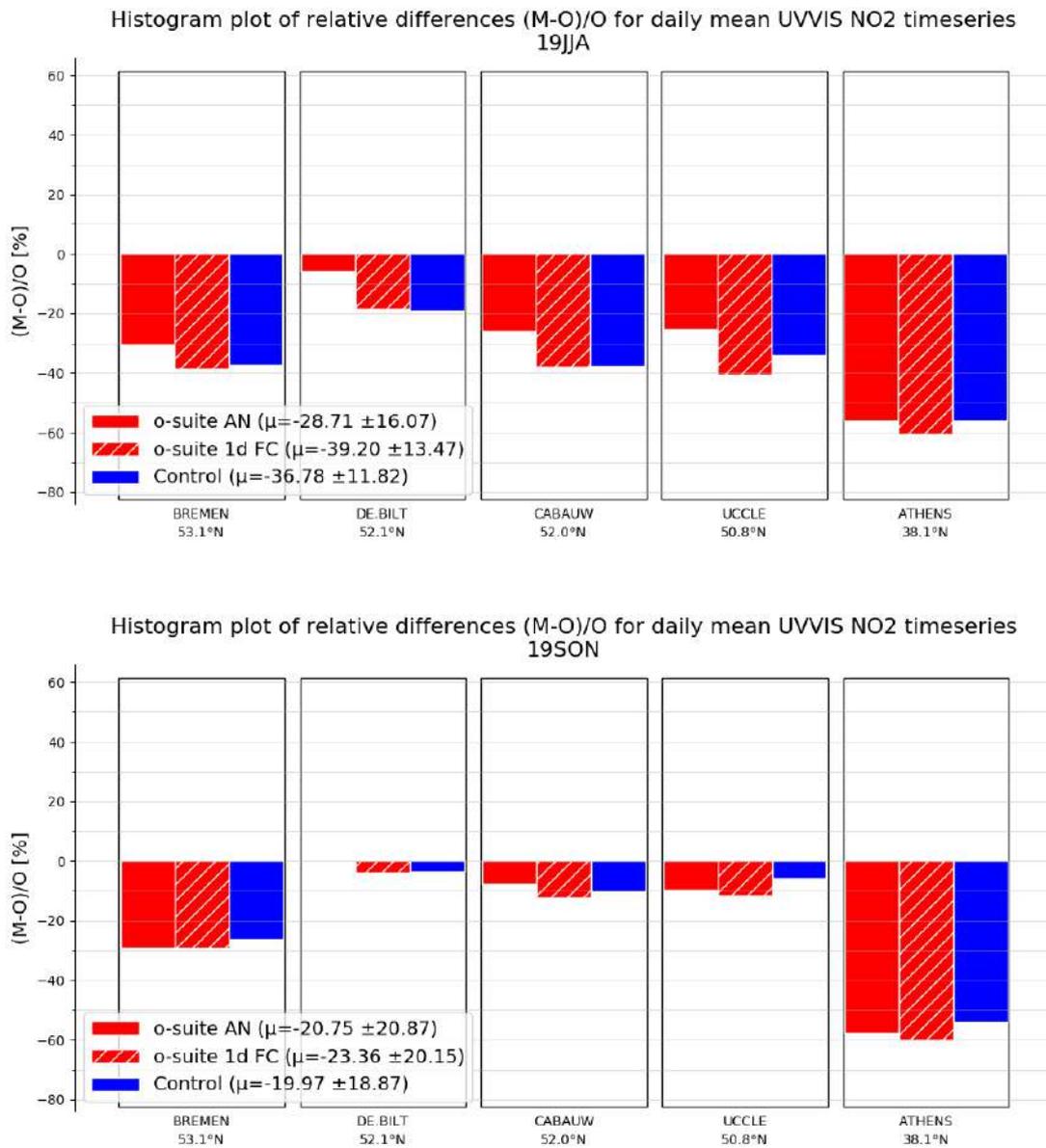
Figure 5.2.1: Table diagram showing the seasonal bias June-August 2019 (top) and Sep-Nov 2019 (bottom), for five stations, sorted by latitude. Compared to the previous validation period JJA, the relative biases in SON did not change significantly.
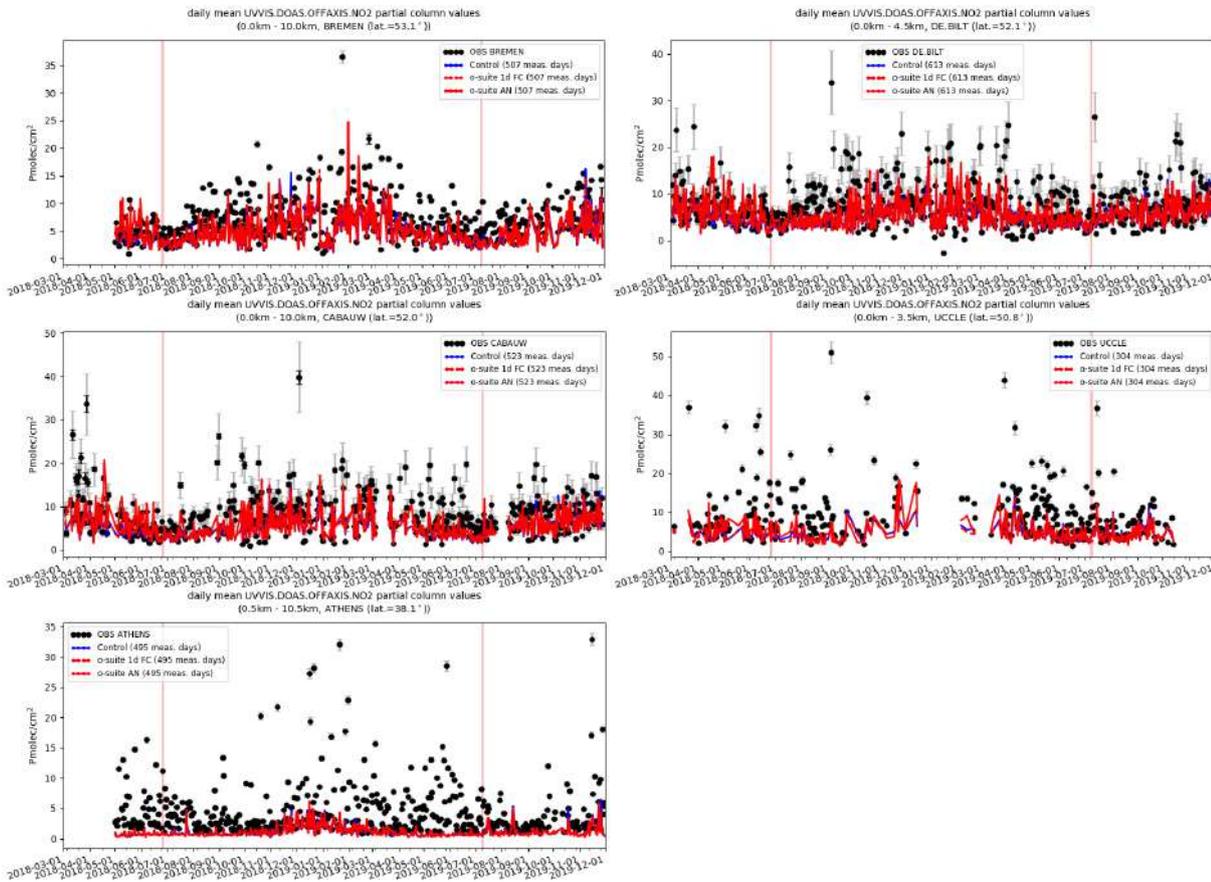
Figure 5.2.2: Time series of NO₂ partial columns at the five different sites. For all sites except Athens, background concentrations are well captured by the CAMS products. The o-suite and control product show little difference.

6. Formaldehyde

6.1 Validation against satellite data

In this section, simulations of tropospheric formaldehyde are compared to SCIAMACHY/Envisat HCHO satellite retrievals (IUP-UB v1.0) [Wittrock et al., 2006] for model data before April 2012 and to GOME-2/MetOp-A HCHO data (IUP-UB v1.0) [Vrekoussis et al., 2010] afterwards. First comparisons to TROPOMI/Sentinel-5P data (IUP-UB v1.0) are provided, using the CAMS o-suite as a-priori in the TROPOMI retrievals. The HCHO retrievals are described in Alvarado et al. (2019). As the retrieval is performed in the UV part of the spectrum where less light is available and the HCHO absorption signal is smaller than that of NO_2, the uncertainty of monthly mean HCHO columns is relatively large (20% – 40%) and both noise and systematic offsets have an influence on the results. However, absolute values and seasonality are retrieved more accurately over HCHO hotspots.

In Figure 6.1.1, monthly mean satellite HCHO columns from GOME-2 are compared to model results for Sep 2019 as an example for autumn 2019. The magnitude of oceanic and continental background values and the overall spatial distribution are well represented by the o-suite and control. The models regularly overestimate values over regions in Central Africa, which could be due to fire or biogenic emissions. This appears less pronounced for autumn 2019 compared to recent years. Moreover, HCHO columns over regions with fire and biogenic emissions in Northern Australia were regularly overestimated mainly during SON and DJF, but this appears much less pronounced for autumn 2019. The overestimation for tropospheric NO_2 over Alaska reported in section 5.1 does not show up in the HCHO comparisons, showing that the ability of the simulations to reproduce these values varies depending on the species, location and time of the year. So far similar conclusions arise from TROPOMI based map comparisons (see Figure 5.1.2 for Sep 2019) though the magnitude of the differences between observations and simulations differs. Differences in comparison results are in principal due to differences in observation time or differences in the retrieval products.

Time series in Fig. 6.1.3 highlight three cases:

- East-Asia and the Eastern US, where HCHO is dominated by biogenic emissions. Model results and measurements generally agree rather well. However, all model runs underestimate the yearly cycle over East-Asia since 2012. In contrast to MOZART runs, MACC_CIFS_TM5 overestimated satellite values for the Eastern US since the middle of 2013. However, the newer IFS-CB05 runs perform well for Eastern US since 2015. For recent years and both regions, there is virtually no difference between the most recent o-suite run with IFS-CB05 chemistry and the corresponding control run without data assimilation. The variability or "ups and downs" in HCHO columns observed by GOME-2 since December 2014 is due to the lack of data (caused by instrument degradation) for these regions during winter in the Northern Hemisphere, leading to e.g. the negative values in the GOME-2 time series for Eastern US since December 2015. Summertime maxima are still underestimated over East-Asia despite of the higher resolution of the model runs since 2016.
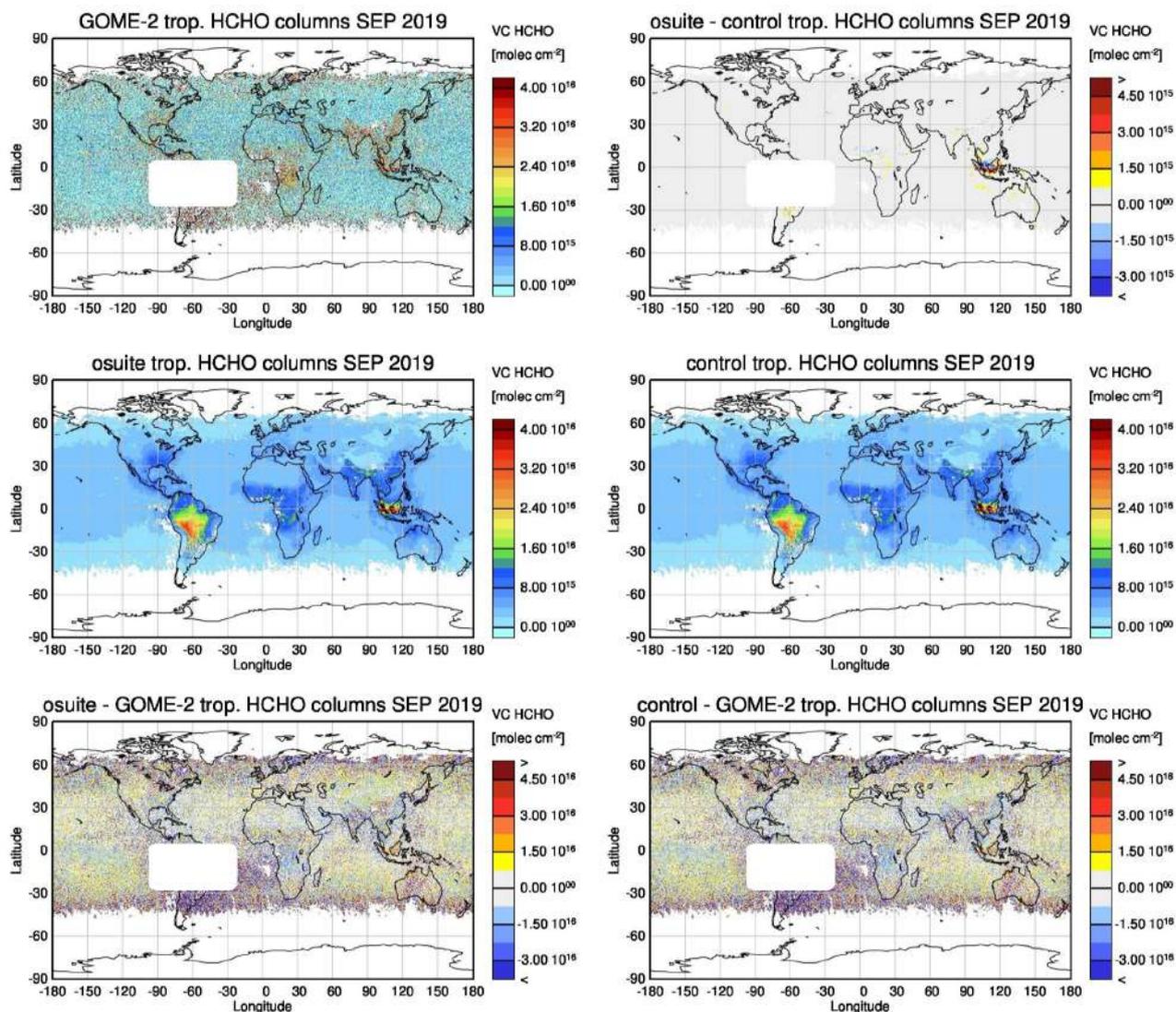
Figure 6.1.1: Global map comparisons of satellite-retrieved and model-simulated tropospheric HCHO columns [molec cm-2] for Sep 2019. The top row shows monthly mean tropospheric HCHO columns retrieved by GOME-2, the second row shows the same but for model simulated averages. The third row shows differences of monthly means between models and GOME-2. GOME-2 data were gridded to model resolution (i.e. 0.4° deg x 0.4° deg). Model data were treated with the same reference sector subtraction approach as the satellite data. Satellite retrieved values in the region of the South Atlantic anomaly are not valid and therefore masked out (white boxes in all images except those which show model results only).

- North-Africa, where biomass burning as well as biogenic sources largely contribute to HCHO and its precursors. Satellite observations over North-Africa tend to be slightly overestimated by IFS-CB05 chemistry model runs since 2014 and also the latest higher resolution model versions since July 2016. However, GOME-2 values are higher, and model values a bit lower for summer 2019 compared to previous years, resulting in a pronounced underestimation with respect to the satellite observations. Moreover, the model simulated HCHO columns increase from summer throughout autumn 2019, though the satellite observed columns show opposite development.
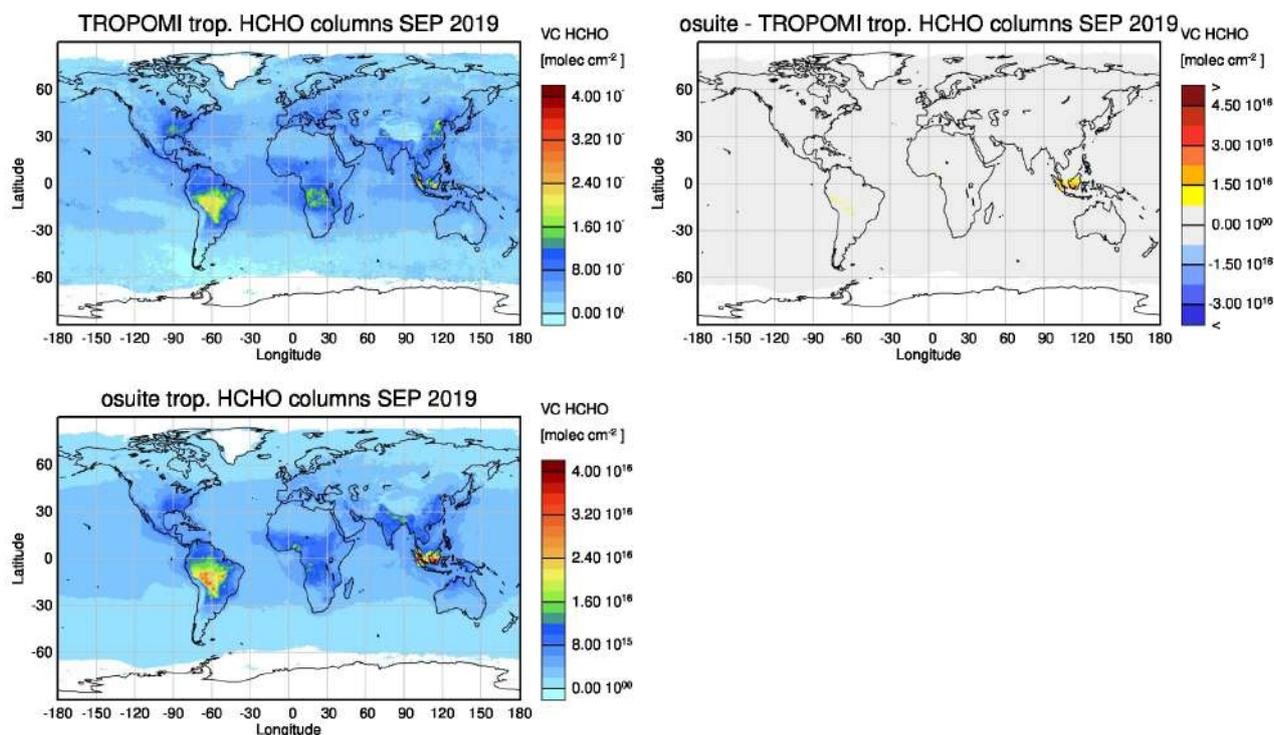
Figure 6.1.2: Global map comparisons of satellite retrieved, and model simulated tropospheric HCHO columns [molec cm^{-2}] for Sep 2019: (top left) TROPOMI, (top right) o-suite minus TROPOMI, (bottom left) o-suite. TROPOMI data were gridded to model resolution (i.e. 0.4° x 0.4° degree) and the CAMS o-suite was used as a-priori in the retrievals. Comparisons to the control are not available for this report.

- Indonesia, where HCHO is also dominated by biogenic sources and biomass burning. Old MOZART based model versions generally overestimated satellite values here (by a factor of 3 – 4 in the second half of 2010) and failed to reproduce the observed seasonality. This may be due to the use of fire emissions including El Nino years, which experience much larger fire activities. MOZART simulations and observations agreed much better since late 2012. IFS-CB05 runs agree very well with satellite retrieved ones for December 2014 to August 2015. For September and October 2015, satellite retrieved HCHO columns show a pronounced maximum. 2015 was a strong El Nino year, which caused droughts and higher fire activity in Indonesia. Another pronounced, but by the models overestimated, increase in satellite observed values associated with comparatively weaker El Nino conditions occurs for Sep 2019. As for previous El Nino years, fire emissions used by IFS-CB05 seem to be largely overestimated, resulting in model-simulated HCHO columns, which are up to twice as large as those retrieved by GOME-2. Further investigations (see previous reports) show that this is not caused by cloud flagging applied to the satellite and model data. Between the middle of 2016 and Sep 2019 there was mainly little variation from one month to another in both, satellite observations and model simulations and the magnitude of model and satellite values agreed overall well.
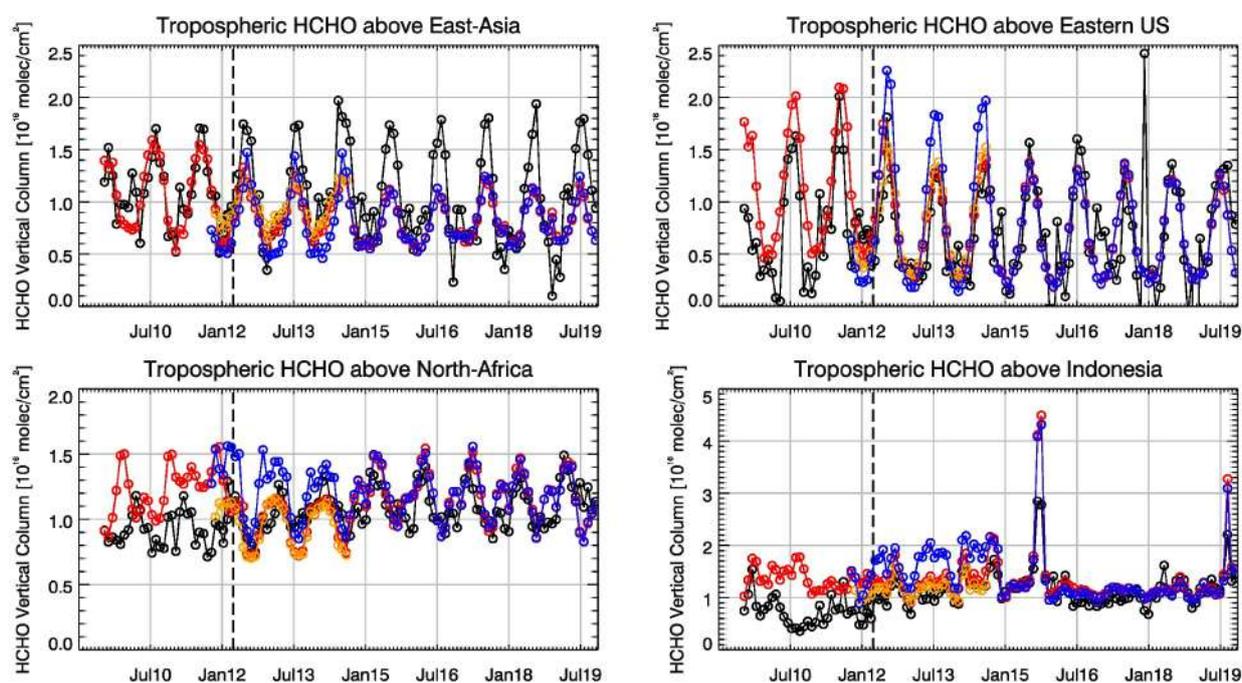
Figure 6.1.3: Time series of average tropospheric HCHO columns [10^{16} molec cm-2] from SCIAMACHY (up to March 2012, black) and GOME-2 (from April 2012 onwards, black) compared to model results (red - osuite, blue - MACC_fncrt_TM5/MACC_CIFS_TM5/control, orange - MACC_fcnrt_MOZ) for different regions. The blue line shows MACC_fcnrt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results from December 2012 to August 2014 and control results from September 2014 onwards (the model run without data assimilation is termed control since Sep 2014). The regions differ from those used for NO_2 to better focus on HCHO hotspots: East-Asia (25-40°N, 110-125°E), Eastern US (30-40°N, 75-90°W), Northern Africa (0-15°N, 15°W-25°E) and Indonesia (5°S-5°N, 100-120°E). Negative satellite retrieved values over Eastern US are due to a lack of data (caused by instrument degradation) during Northern Hemisphere winter months for this region. Vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

Time series comparisons between the o-suite and TROPOMI are shown in Figure 6.1.4 since September 2019. They show differences with respect to the GOME-2 based ones: the peak over Indonesia for Sep 2019 is much less pronounced for both the o-suite and satellite observations, the development of values in time is in agreement over North-Africa. Apart from Indonesia, differences between observations and the o-suite are generally less pronounced. Differences in comparison results are in principal due to differences in observation time or differences in the retrieval products.

For details on the HCHO evaluation:
http://www.doas-bremen.de/macc/macc_veri_iup_home.html
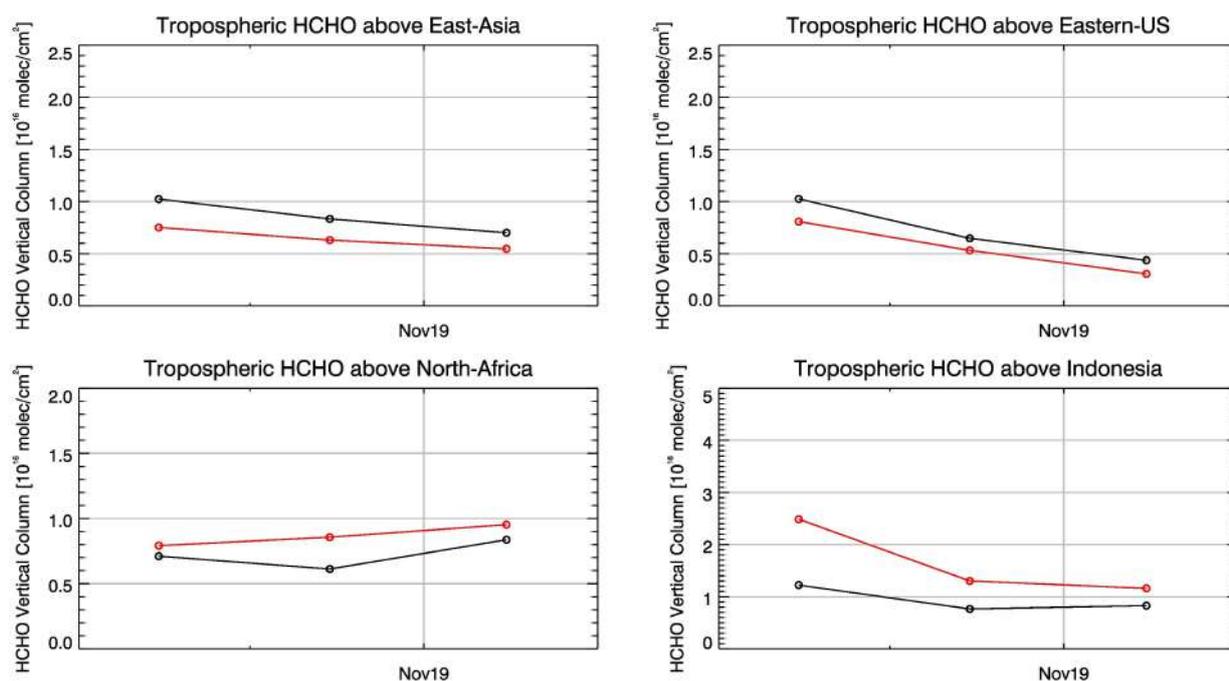
Figure 6.1.4: Time series of average tropospheric HCHO columns [10^{16} molec cm^{-2}] from (black) TROPOMI compared to (red) o-suite model results since September 2019 (see Annex 2 for definition of regions). The regions differ from those used for NO_2 to better focus on HCHO hotspots: East-Asia (25-40°N, 110-125°E), Eastern US (30-40°N, 75-90°W), Northern Africa (0-15°N, 15°W-25°E) and Indonesia (5°S-5°N, 100-120°E).

6.2 Evaluation against ground-based DOAS observations

In this section, we compare the HCHO columns of the CAMS products with UVVIS DOAS measurements at Uccle, Cabauw and De Bilt.[2] These ground-based, remote-sensing instruments are sensitive to the HCHO abundance in the lower troposphere. Tropospheric HCHO profiles and columns are validated (up to 3.5km (Uccle) or 10km (Cabauw and De Bilt)). The validation methodology is the same as for the MWR O_3 and FTIR O_3 and CO validations see http://nors.aeronomie.be. It is important to mention here that the model partial column values are calculated for the smoothed model profiles. This guarantees that the model levels where the measurement is not sensitive do not contribute to the observed bias. We should mention that the measurement data is catalogued as rapid delivery and not in the consolidated NDACC database.

Fig. 6.2.1 shows the absolute biases September – November 2019 at the different sites and indicates nearly vanishing bias for the different sites. At Cabauw, some high pollution events are not captured by the model and leads to a higher overall underestimation (Fig 6.2.2). From Fig. 6.2.1 and 6.2.2 we see little difference between the o-suite and the control run. Although the background column values are well captured by the products, the high emission events are not.

---

[2] No contribution from Reunion, Xianghe and OHP due to instrument failure.
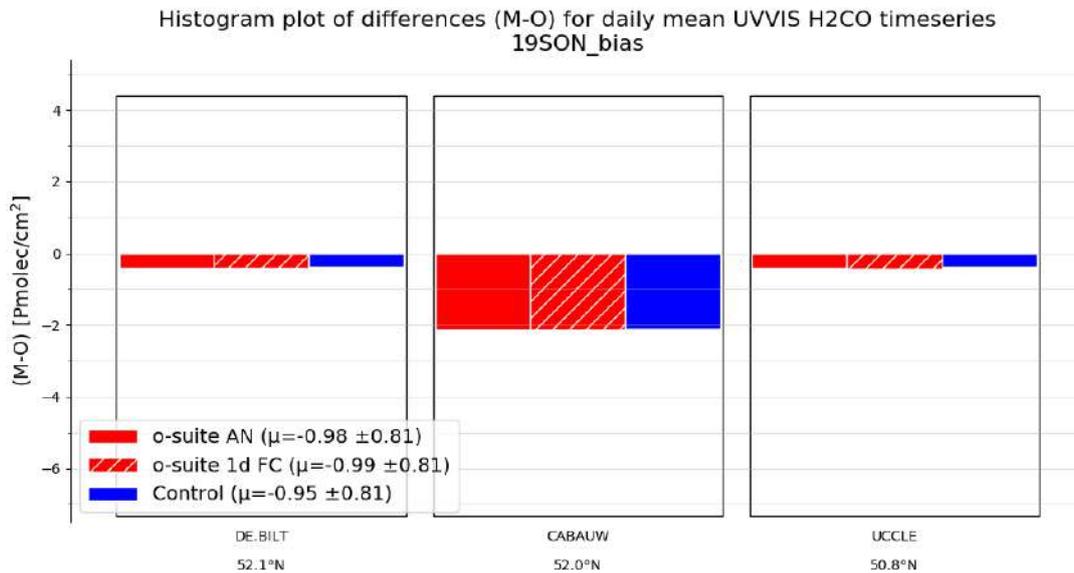
Figure 6.2.1: Table diagram showing the seasonal absolute bias in SON 2019 for three stations, sorted by latitude.
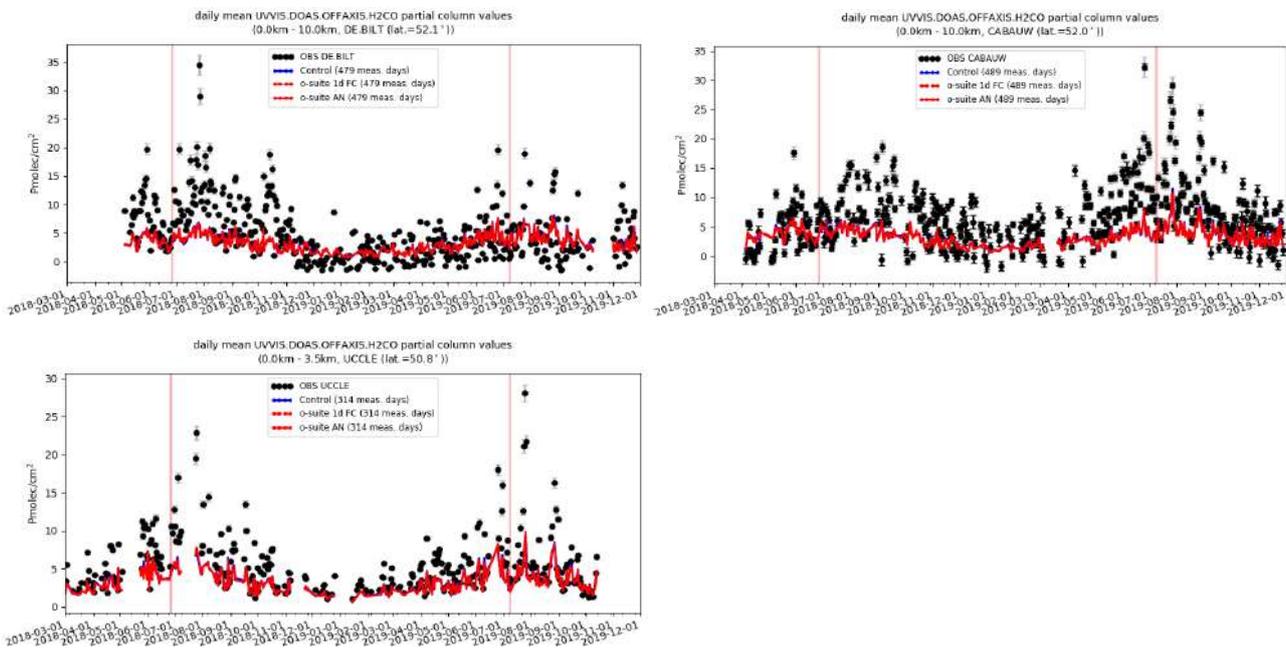
Figure 6.2.2: Time series of H_2CO partial columns at the five different sites. All CAMS products underestimate the H_2CO concentrations, except at De Bilt, where the model overestimates during the winter months.

7. Aerosol

7.1 Global comparisons with Aeronet and EMEP

The comparison of the CAMS simulation of time series of aerosol optical depth can be found for all Aeronet stations at: http://aerocom.met.no/cams-aerocom-evaluation/

More detailed evaluation including scores, maps, scatterplots, bias maps and histograms illustrating the performance of the aerosol simulation in the IFS system are made available through the AeroCom web interface. The model run can be compared here to e.g. the CAMS interim reanalysis and other models, such as the AeroCom Median model.

Correlation, based on daily aerosol optical depth and NRT Aeronet observations, has been rather stable recently. The o-suite forecast at +3 days shows only slightly lower correlation. See figure S3.

Part of the month-to-month variation in correlation is due to the varying quality and coverage of the Aeronet network. This has been improved by the version 3 from Aeronet. We use therefore version 3 level 1.5 for all global comparison to Aeronet.

 The performance of the o-suite model exhibits some seasonal variation in AOD depending on region (Fig. 7.1.1). Noteworthy is the persistent AOD overestimation over North America (Fig. 7.1.1-bottom), but also a long-term trend to overestimation in East Asia. The latitudinal display of model and Aeronet AOD in the period investigated here (Fig. 7.1.2) shows a small positive bias against Aeronet in the Southern Hemisphere.
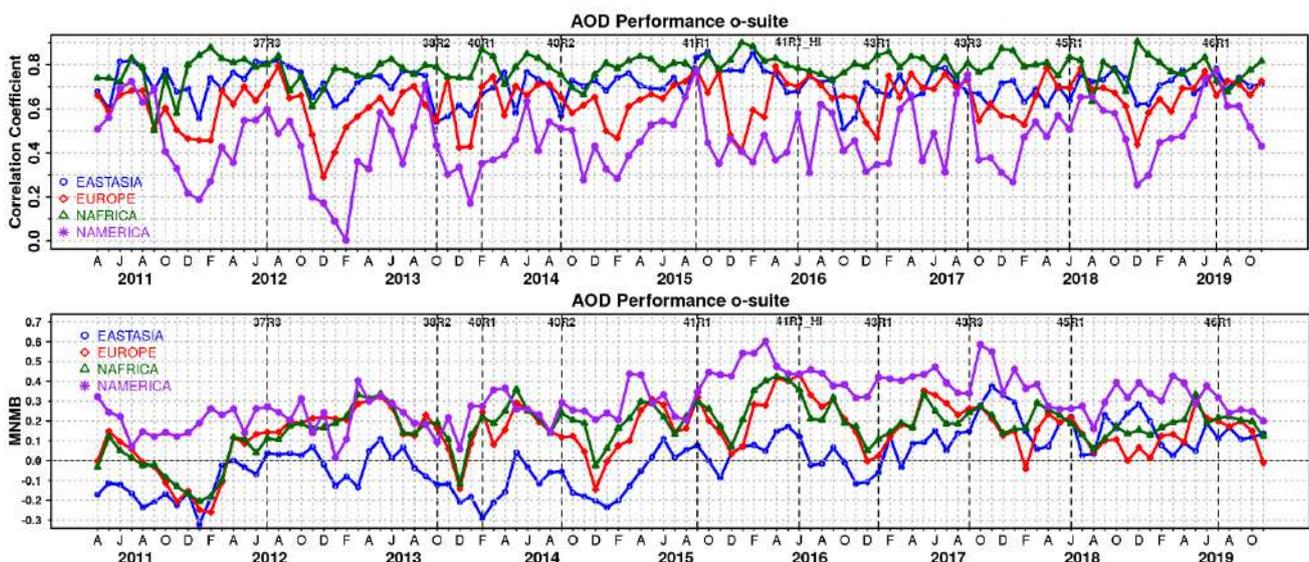
Figure 7.1.1. (top) Correlation coefficient and (bottom) modified normalized mean bias (MNMB) in AOD, since 2011, based on daily AOD comparison (Aeronet V3 level 1.5 data) in four world regions [East-Asia (blue); Europe (red); North Africa (green); North America (purple)] for the o-suite.
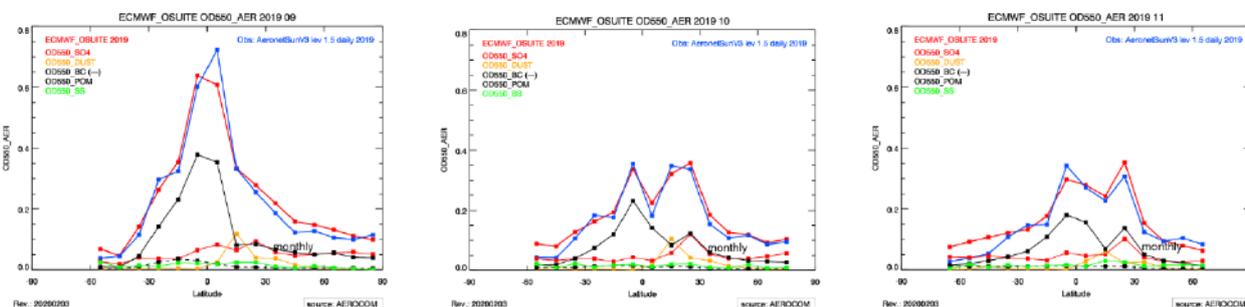
Figure 7.1.2. Aerosol optical depth of o-suite (red) compared to latitudinally aggregated Aeronet V3 level 1.5 data (blue) for the three months covered by this report.
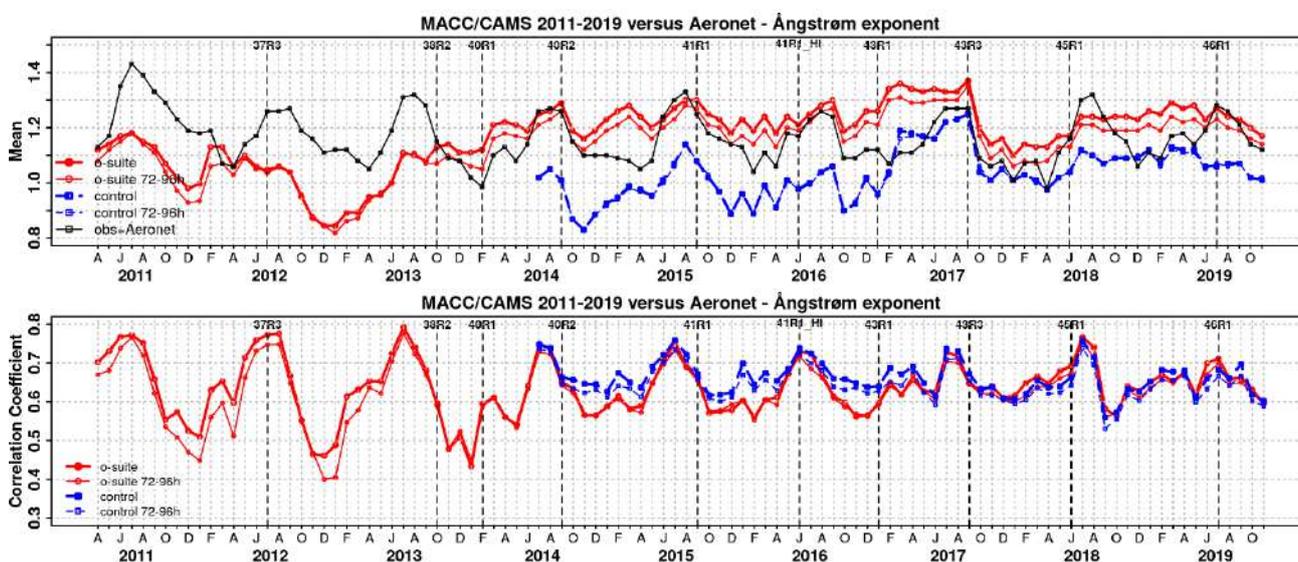
Figure 7.1.3. a) (top) Evolution of mean Ångström exponent in o-suite and control at Aeronet sites (Aeronet V3 level 1.5 data), based on matching monthly mean values. o-suite (thick red curve); o-suite at last forecast day (light red curve); control (blue dashed curve); control at last forecast day (light blue dashed curve). b) (bottom) Correlation using daily matching Ångström exponent.

o-suite				
	Mean JJA 2019 0-24h	Change wrt to first day on day 4	Mean SON 2019 0-24h	Change wrt to first day on day 4
AOD@550	0.181	-11%	0.158	-14%
BC-OD@550	0.006	-16%	0.005	-19%
Dust-OD@550	0.028	10%	0.015	13%
OA-OD@550	0.048	-23%	0.048	-24%
SO4-OD@550	0.043	-22%	0.036	-22%
SS-OD@550	0.053	-3%	0.053	-6%

Table 7.1.1. Mean global total and speciated AOD in the o-suite for the last two periods covered by the VAL report and change after 3 forecast days.
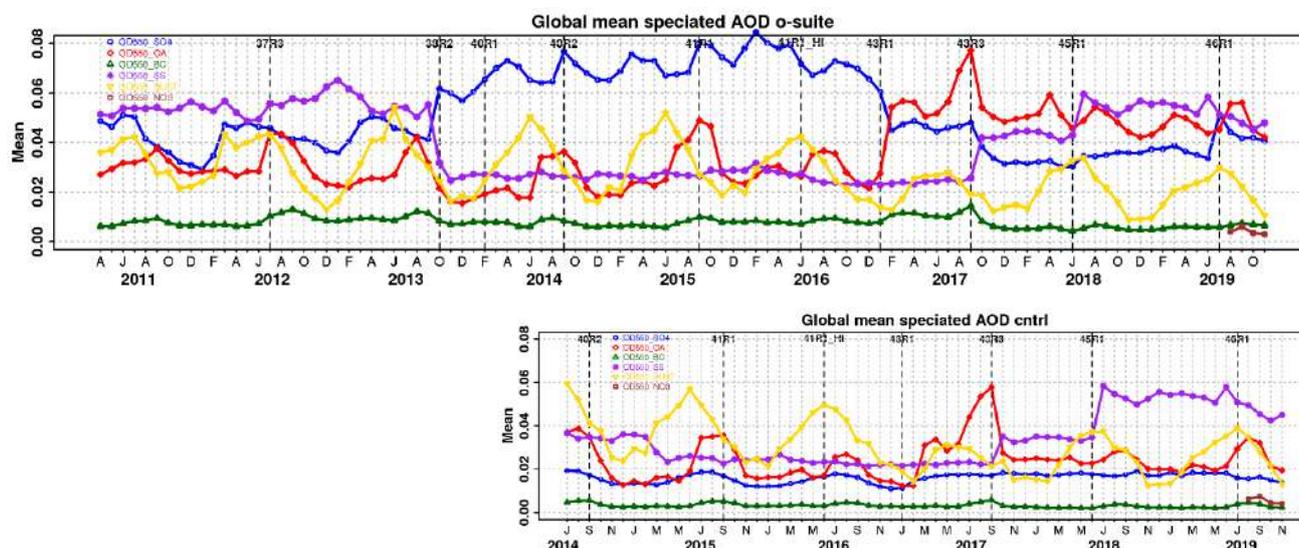
Figure 7.1.4. Evolution of the aerosol components of total AOD@550nm [OD550_SO4 = sulphate(blue); OD550_OA = organics(red); OD550_BC = black carbon(green); OD550_SS = sea salt(purple); OD550_DUST = dust(yellow); OD550_NO3 = nitrate(brown)] in o-suite and control simulation.

The simulated aerosol size distribution may be validated to first order using the wavelength-dependent variation in AOD, computed as Ångström exponent, with higher Ångström exponents indicative of smaller particles. We find in SON 2019 a small bias (Figure 7.1.3-a). Temporal and spatial variability is difficult to capture, but correlation from all daily data is lower than for AOD (Figure 7.1.3-b and S3). Figure 7.1.4 shows that the Sep 2017 and Jun 2018 model changes are responsible for a shift in Ångström exponent. More organic matter seems to shift the size distribution to smaller sizes. The model upgrade in Feb 2017 with a bugfix for sea salt and improved parameterisations for SO4 lead to that sea salt increased with 45% while sulphate further decreased a bit. Sea salt has increased further due to a new sea salt emission scheme implemented in the Jun 2018 model upgrade and is back to earlier 2011-2013 levels. Since the latest model upgrade with the improvement to the sulphur cycle, the SO4 seem to have increased to same levels as before the Sep 2017 upgrade.

The o-suite uses data assimilation to obtain an analysis of the aerosol field. In the forecast period, however, a-priori model parameterisations and emissions (except fire emissions, which are kept in the forecast equal to the latest GFAS emission values) determine increasingly the aerosol fields. The performance of the day three forecasted AOD fields as compared to the first guess is shown in Figure S3 in the summary of this report. Table 7.1.1 shows an average global decrease in total aerosol optical depth during the first four forecast days, dominated by sulphate and organics. The control run with no assimilation shows somewhat less AOD (-31% compared to o-suite, see figure S3). All this supports the conclusion that either a-priori IFS aerosol and aerosol precursor sources are too small, or sinks are too effective in the IFS model.
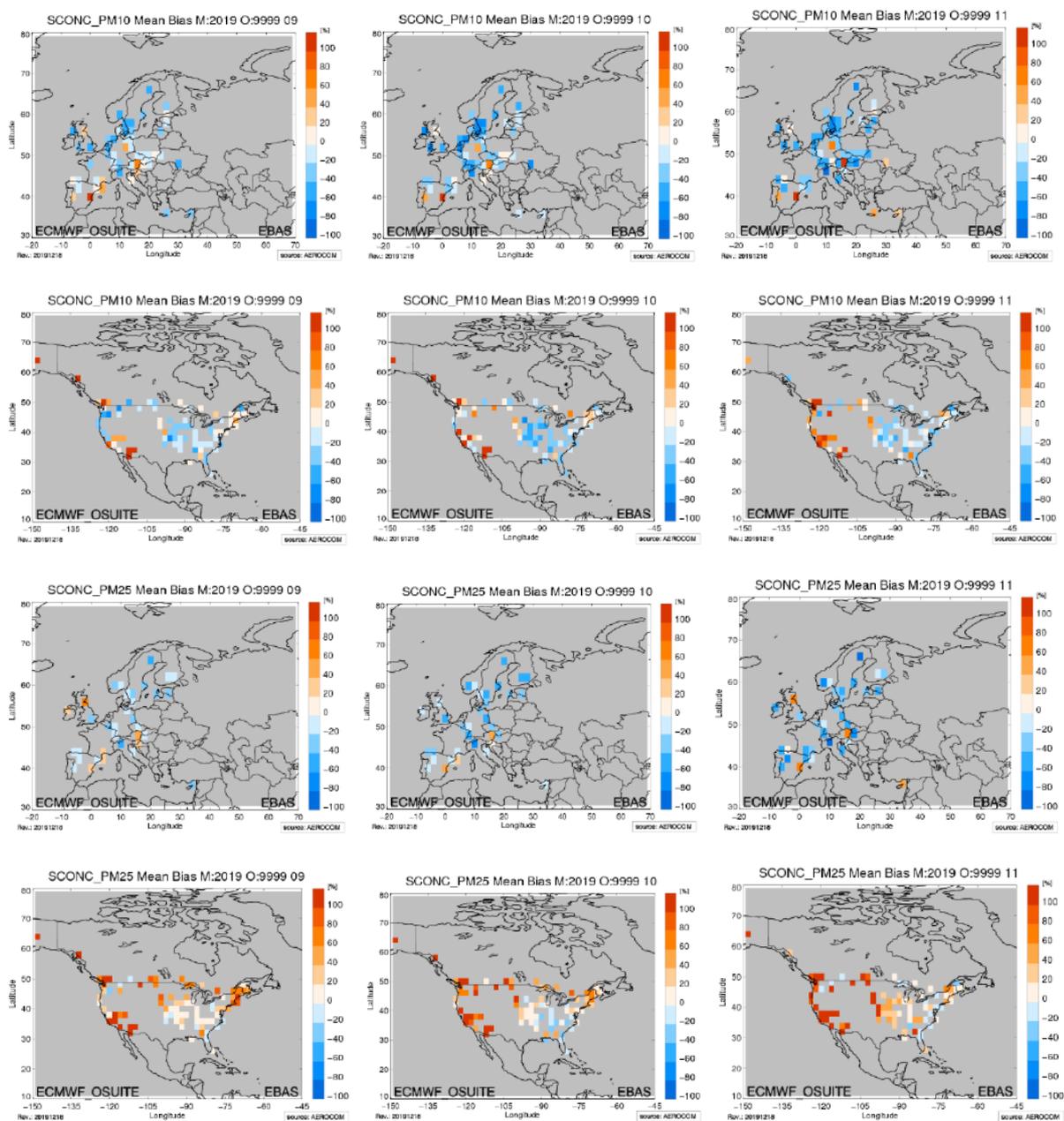
Figure 7.1.5. Bias [%] map of monthly mean PM10 and PM2.5 concentrations at EMEP (Europe, first and third row) and IMPROVE sites (North America, second and fourth row) for September (left column), October (middle) and November 2019 (right); simulated o-suite versus EMEP/IMPROVE derived climatological average (2000-2009).
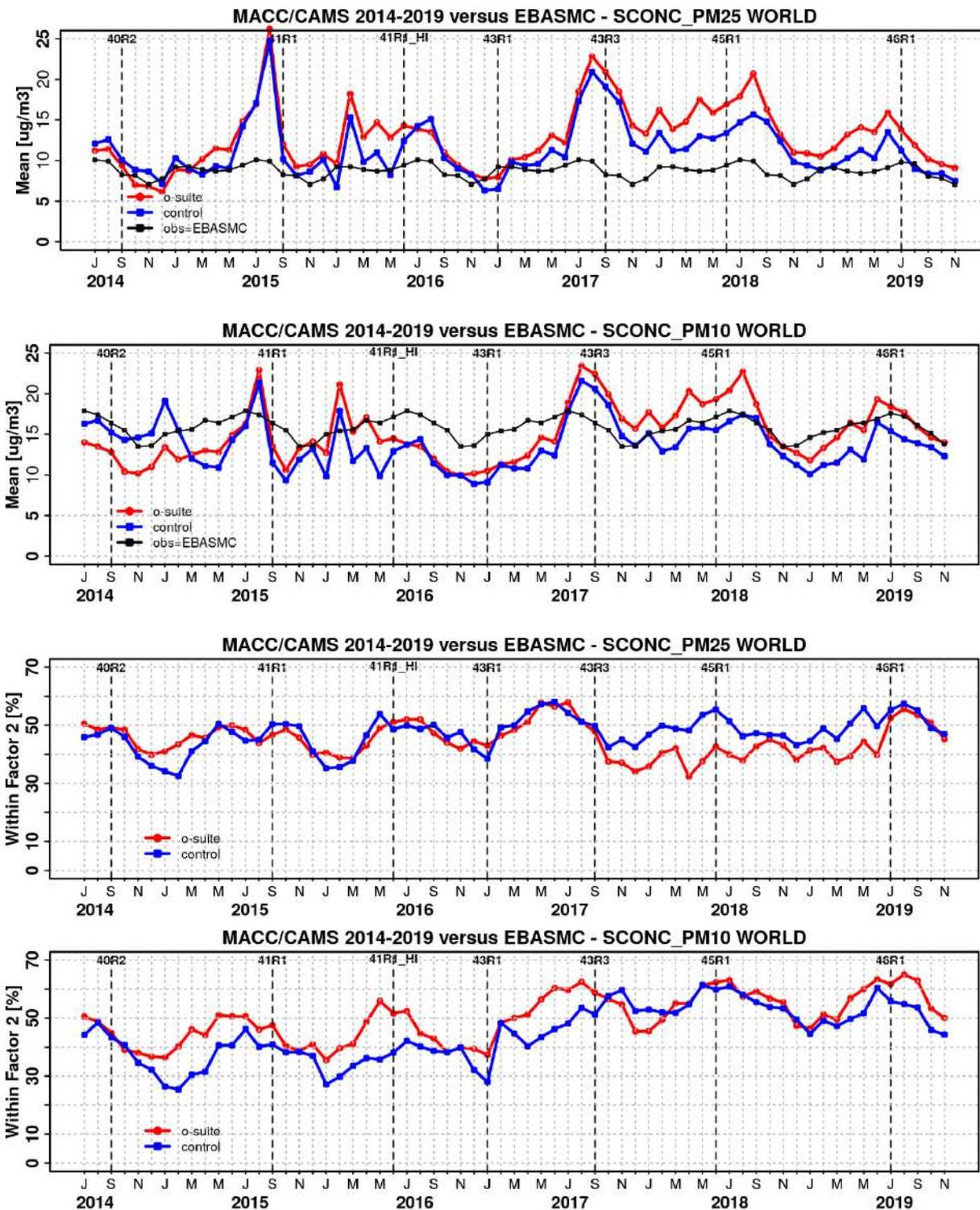
Figure 7.1.6. Temporal evolution of monthly mean average PM25 and PM 10 concentrations at EMEP (Europe) and IMPROVE sites (North America) and data fraction within a factor 2 of observed; ca 160 sites, observed data averaged from data available in EBAS from 2000-2009.

Surface concentration of particulate matter below 10 µm (PM10) and below 2.5 µm (PM25) from the o-suite experiment have been validated against data from 160 background IMPROVE and EMEP stations. A climatological average has been constructed from data in the period 2000-2009 as available in the EBAS database hold at NILU. The data availability is not the same at all stations, and sometimes covers only a few years.

A negative MNMB bias of PM10 in Europe and an overestimate in North America PM2.5 appears (Fig. 7.1.5), consistent with the AOD bias in the two regions. Figure 7.1.6 shows the evolution of mean observed and simulated PM10 and PM2.5. The biggest change appeared in July 2017 with the bias of o-suite now becoming positive almost overall. Shown is also the statistics of being within factor 2, a more robust metrics for a comparison to climatological data. This statistical indicator has clearly improved over time, indicating best PM10 and PM25 performance in summer months for the o-suite. O-suite is also better most of the times than the control simulation for PM10. For PM25 the difference is less clear, but since September 2017 (upgrade to 43R3) the control is performing better than the o-suite, except since latest upgrade with a performance of o-suite as good as the control.

7.2 Validation of dust optical depth against AERONET, and comparisons with the Multi-model Median from SDS-WAS

The 72-hour forecasts (on a 3-hourly basis) of dust aerosol optical depth (DOD) from CAMS o-suite and control have been validated for the period 1 September 2019 – 30 November 2019 against the AERONET Spectral Deconvolution Algorithm (SDA) cloud-screened observations, MODIS/Terra and Aqua Collection 6.1 Level 3 (1º x 1º) and SDS-WAS Multi-model Median DOD. The SDS-WAS Multi-model Median DOD is obtained from (currently) twelve dust prediction models participating in the Sand and Dust Storm Warning Advisory and Assessment System (SDS-WAS) Regional Center for Northern Africa, Middle East and Europe (http://sds-was.aemet.es/). At those sites where the SDA products are available, the dust AOD evaluation will be complemented with AOD-coarse, which is fundamentally associated to maritime/oceanic aerosols and desert dust. Since sea-salt is related to low AOD (< 0.03; Dubovik et al., 2002) and mainly affects coastal stations, high AOD-coarse values are mostly related to mineral dust.

During this season, satellites (see MODIS in Figure 7.2.1) show that major dust activity in Northern Africa (seasonal AOD up to 0.3) is concentrated in latitudes between 15 and 20ºN with maximum seasonal AOD values over 0.7. Meanwhile, in the Middle East, high AOD values up to 0.3 are observed in the Iraq and Saudi Arabia. In general, the CAMS o-suite shows lower season values than the control run, which are in general higher than the SDS-WAS multi-median product.

The CAMS o-suite can reproduce the dust transport over the North Atlantic region, although the maximum dust activity is shifted to Mauritania, Mali, Niger and Algeria border and Sudan, in addition to Chad as shown in MODIS (see Figure 7.2.1). Also, DOD over Iraq and in the Mediterranean Basin appears overestimated in comparison with the SDS-WAS multi-model ensemble. These changes in dust activity in the main source regions are linked to the new dust module implemented in the operational CAMS model since early-July 2019.
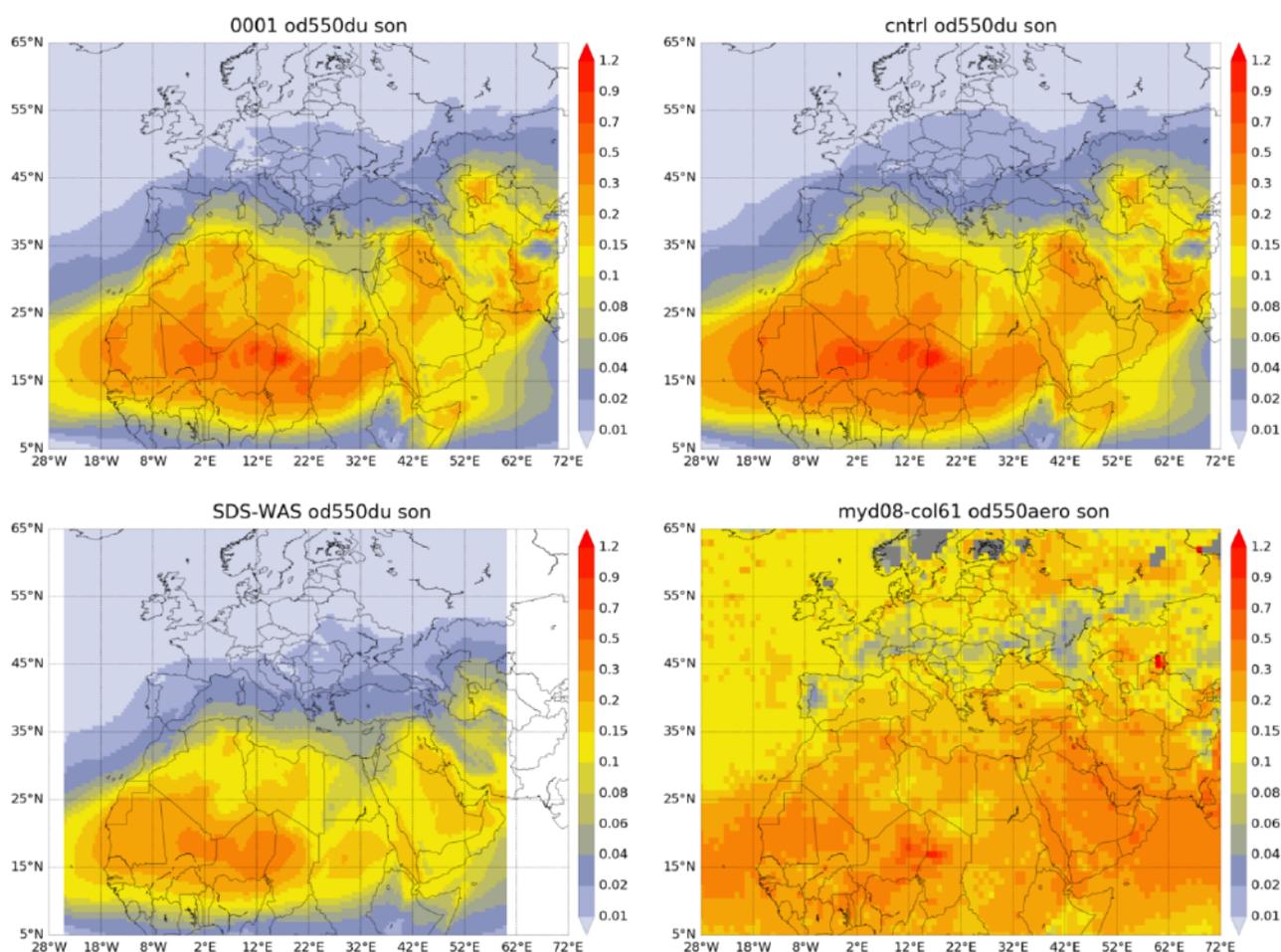
Figure 7.2.1: Averaged DOD 24h forecast from o-suite (top left) and control (top right), DOD of the multi-model SDS-WAS Median product (bottom left) as well as AOD from MODIS/Aqua Collection 6.1 Level 3 combined Dark target and Deep Blue product (bottom right) for the study period.

From September to November, the o-suite and control experiments reproduce the daily variability of AERONET direct-sun observations (see Figure 7.2.2a and Table 7.2.1), with a correlation coefficient of 0.77 averaged over all AERONET sites, which is lower than the SDS-WAS multi-model product which has a correlation coefficient of 0.82. Regarding mean bias (MB), the CAMS o-suite tends to reduce the overestimations observed in control with a MB of -0.01 for o-suite and 0.03 for the control experiment. The SDS-WAS multi-model underestimates (MB of -0.03) the AERONET dust-filtered observations. Similar results are obtained in comparison with the AERONET SDA observations (see Figure 7.2.2b).

Over desert dust sources in the Sahara (see Table 7.2.1 as well as Tamanrasset INM AERONET site in Figure 7.2.3a), CAMS does reproduce the daily variability with correlation coefficients of 0.68 and 0.72 for o-suite and control. Overestimations are observed in both CAMS products over the Sahara (MB of 0.15 for control and 0.06 for o-suite). The SDS-WAS Multi-model results for Sahara shows better skills for this season (with a seasonal correlation of 0.77 and MB of -0.02). In the Middle East, the comparison with AERONET direct-sun dust-filtered observations shows correlations
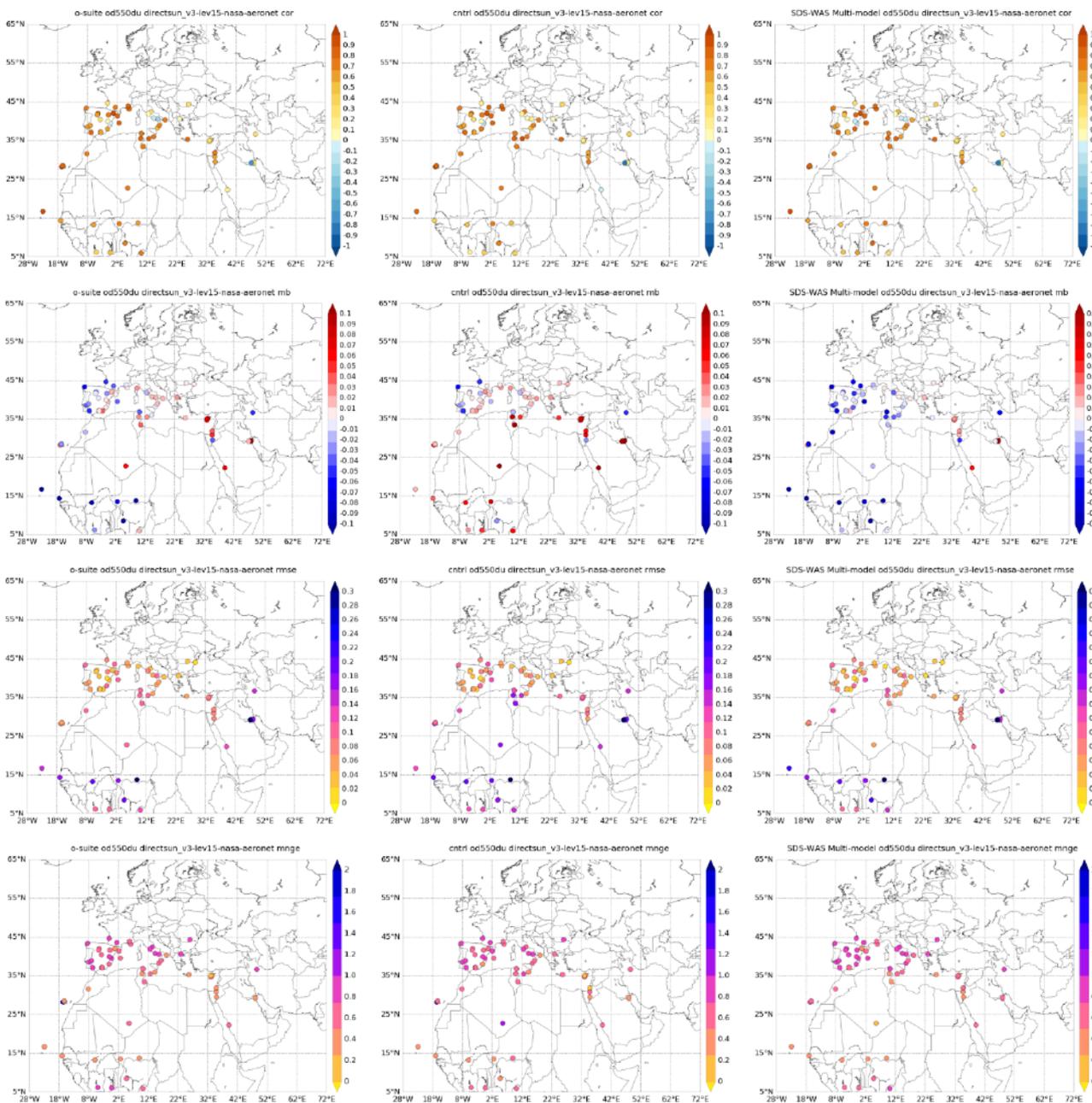
Figure 7.2.2a: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite (left column), control (central column) and DOD Multi-model SDS-WAS Median (right column) for the study period. Dust-filtered AOD from AERONET direct-sun is the reference.

around 0 (see Table 7.2.1), due to the low number of observations available. In the case of the AERONET SDA AOD coarse product (see Figure 7.2.2b), correlations are 0.53 for control and 0.59 for o-suite. This difference between both datasets is associated with the lower AOD concentrations during this season and the presence of smaller aerosols (see AE > 1.2 in Kuwait University AERONET site in Figure 7.2.3a).
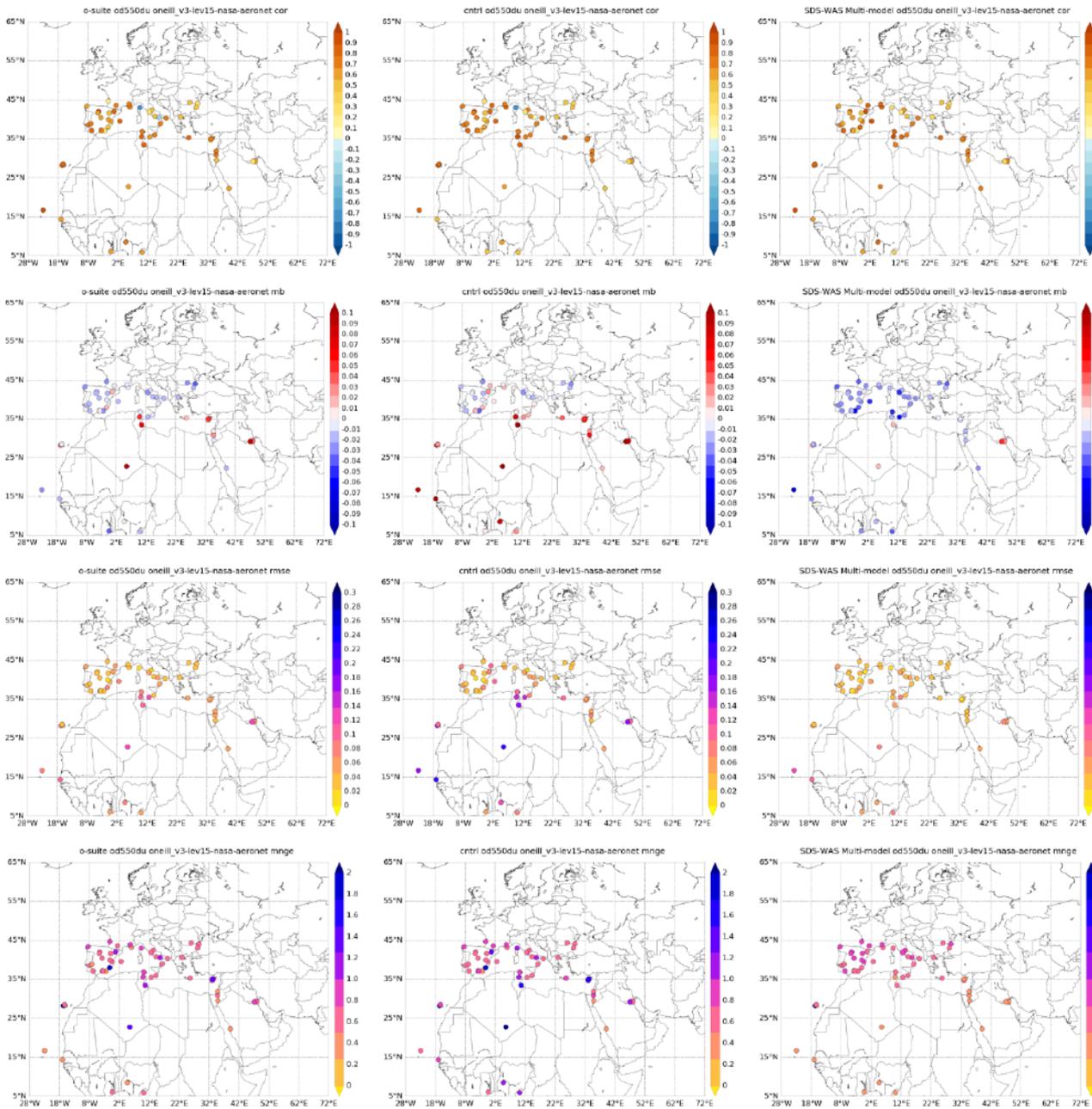
Figure 7.2.2b: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite (left column), control (central column) and DOD Multi-model SDS-WAS Median (right column) for the study period. AOD-coarse from AERONET SDA is the reference.

Table 7.2.1: Skill scores (MB, FGE, RMSE and r) of 24h forecasts for CAMS o-suite, CAMS control and SDS-WAS Multi-model Median for the study period, and the number of data (NDATA) used. Dust AOD (DOD) from AERONET is the reference.

	NDATA	Control				o-suite DOD				SDS-WAS Median DOD			
		MB	FGE	RMSE	r	MB	FGE	RMSE	r	MB	FGE	RMSE	r
Sahara	274	0.15	0.71	0.18	0.72	0.06	0.39	0.11	0.68	-0.02	-0.01	0.07	0.77
Sahel	881	0.04	0.17	0.25	0.57	-0.11	-0.25	0.24	0.64	-0.12	-0.24	0.25	0.70
Tropical North Atlantic	139	0.02	-0.14	0.14	0.86	-0.11	-0.50	0.15	0.90	-0.17	-0.68	0.21	0.92
Subtropical North Atlantic	397	0.03	0.27	0.13	0.71	-0.01	0.07	0.07	0.73	-0.04	-0.09	0.07	0.73
North Western Maghreb	89	0.02	-0.11	0.12	0.69	-0.02	-0.26	0.10	0.69	-0.10	-0.90	0.14	0.68
Western Iberian Peninsula	297	-0.03	-0.10	0.07	0.62	-0.04	-0.13	0.07	0.61	-0.05	-0.24	0.07	0.60
Iberian Peninsula	350	-0.01	0.85	0.06	0.71	-0.02	0.82	0.06	0.70	-0.03	0.68	0.06	0.70
Western Mediterranean	1272	0.02	1.30	0.08	0.79	0.00	1.26	0.07	0.76	-0.02	1.15	0.06	0.82
Central Mediterranean	954	0.04	1.07	0.11	0.81	0.01	1.00	0.08	0.82	-0.02	0.82	0.08	0.79
Eastern Mediterranean	664	0.06	1.70	0.09	0.48	0.05	1.67	0.08	0.48	0.02	1.63	0.06	0.56
Eastern Sahara	-	-	-	-	-	-	-	-	-	-	-	-	-
Middle East	233	0.06	1.41	0.17	-0.02	0.04	1.40	0.15	0.02	0.03	1.37	0.14	-0.04
All sites	**5550**	**0.03**	**0.87**	**0.14**	**0.77**	**-0.01**	**0.73**	**0.12**	**0.77**	**-0.04**	**0.61**	**0.12**	**0.82**

As shown in Figure 7.2.3a, the higher CAMS overestimations of the control run are reduced in the o-suite experiment. As indicated before, these changes in the dust activity in the main source regions are linked to the new dust module implemented in the operational CAMS model since 9th July 2019.

In the Sahel (see Figure 7.2.2 and Table 7.2.1), the o-suite shows strong underestimations (MB of -0.11) in comparison with control (with MB of 0.04). However, the o-suite better reproduces the observed daily variability (with a correlation of 0.64 for o-suite in comparison to 0.57 for control). The underestimations observed in o-suite in the Sahel are also spread to the Tropical North Atlantic (MB of -0.11 for o-suite, see Dakar in Figure 7.2.3b and Table 7.2.1).

Over long-range transport regions (see Figure 7.2.2), the CAMS o-suite model results perform better than control over the sub-Tropical North Atlantic region with correlations of 0.71 for control and 0.73 for the o-suite (see Table 7.2.1).

Over the Iberian Peninsula and the Mediterranean, both CAMS products show high correlations between 0.70 and 0.82 (except in the Eastern Mediterranean) and slight overestimations that are lower for o-suite (MB between 0 and 0.05). The Central Mediterranean region (see Table 7.2.1 and Tunis Carthage in Figure 7.2.3b) has the best results in the AERONET comparison in terms of correlation. Both experiments can reproduce the daily variability with a correlation coefficient of 0.82 and 0.81 for o-suite and control, respectively. The Eastern Mediterranean (see Table 7.2.1 and Ben Salem in Figure 7.2.3b) presents slightly lower correlation values (0.48 for control and
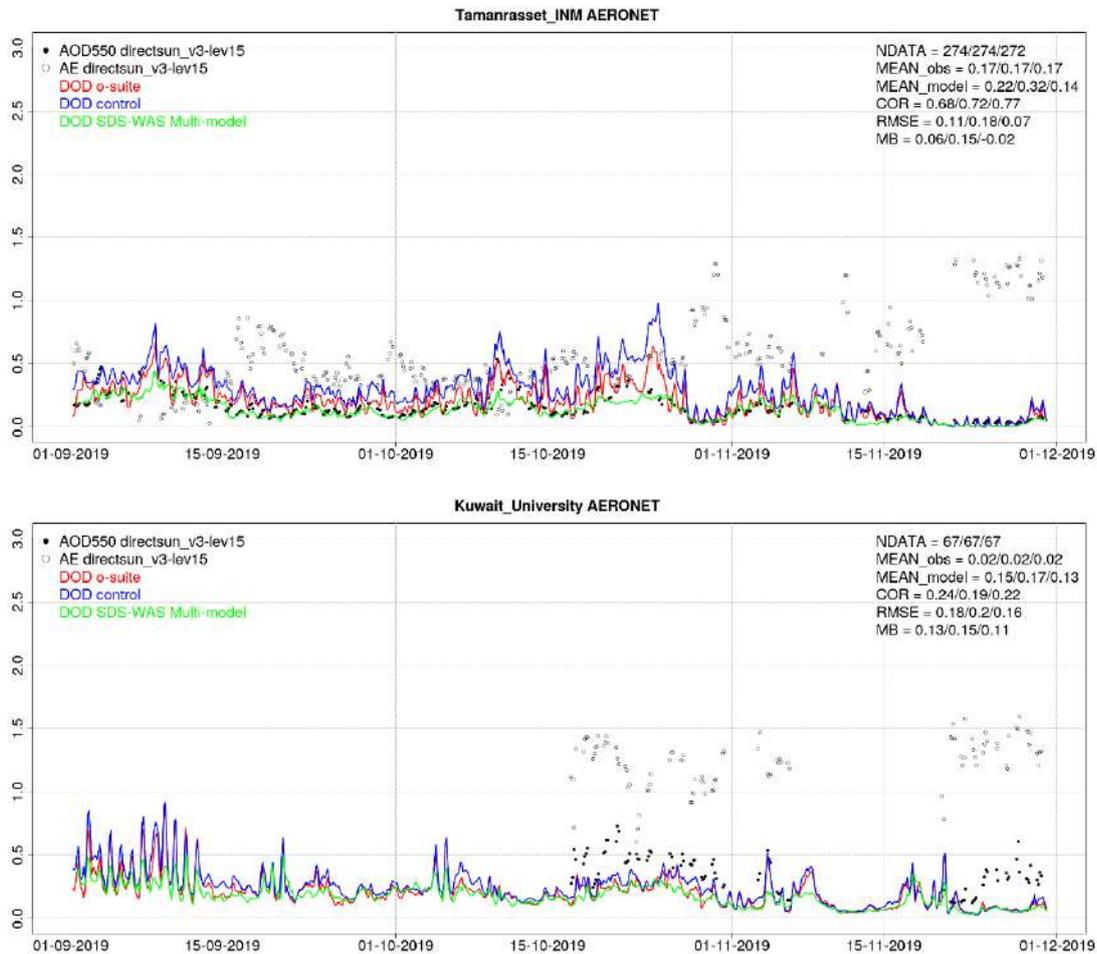
Figure 7.2.3a: AOD and Angstrom Exponent from AERONET Direct-sun (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multi-model SDS-WAS Median (green line) for the study period over Tamanrasset_INM (Sahara) and Kuwait University (Middle East). Skill scores per each individual site and model (o—suite/control/ SDS-WAS Multi-model) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN model, COR, RMSE, MB).

o-suite). This is related to the minimum in dust activity during this period in this region and the associated low AOD. In the case of the North-Western Maghreb, both CAMS products can reproduce the daily variability with a correlation coefficient of 0.69, but control is overestimating the observations (MB of 0.02), while o-suite shows a small underestimation (MB of -0.02).

The comparison of the 1- to 3-day forecasts shows that the prediction is stable during the forecasts in comparison with AERONET direct-sun observations with correlation coefficients of 0.77 (0.77), 0.75 (0.76), and 0.75 (0.75) respectively for 24, 48 and 72h forecasts for all the sites, for o-suite (control).
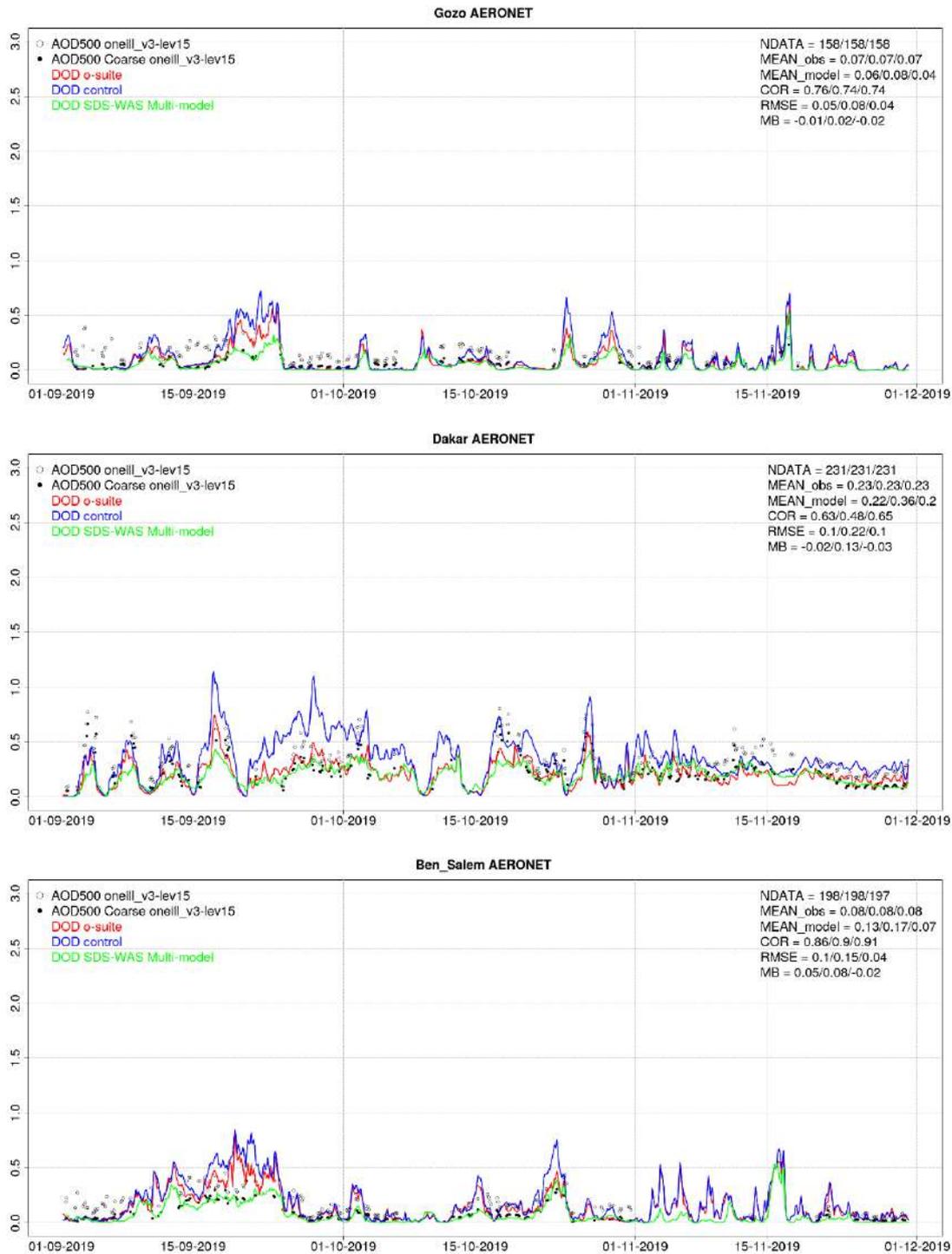
Figure 7.2.3b: AOD and AOD-coarse from AERONET SDA (black dots), DOD o-suite (red line), DOD control (blue line) and DOD Multi-model SDS-WAS Median (green line) for the study period over Gozo (Central Mediterranean), Dakar (Tropical North Atlantic) and Ben Salem (Eastern Mediterranean) Skill scores per each individual site and model (o—suite/control/SDS-WAS Multi-model) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN model, COR, RMSE, MB).

7.3 Ceilometer backscatter profiles

The technical specification of the data sources, the evaluated parameters and methods is provided in the observations report (Eskes et al., 2019). In this section, the temporal and vertical variation of the ceilometer backscatter coefficient (BSC) profiles are evaluated statistically as bias, correlation, and standard deviation of o-suite and control versus ceilometers. The results are summarized in Taylor plots.

The model upgrade of 10 July 2019 increased the number of vertical levels from 60 to 137. Nitrate NO_3 (two size bins) and ammonia NH_4 were included, both coupled to the gas-phase nitrate chemistry, and for sulphur species the chemistry and aerosol schemes have been coupled for consistency. Anthropogenic and biogenic emission inventories, anthropogenic SOA production, biomass burning injection and the dust emission scheme were updated to produce larger particles.

Period Overview

In Figure 7.3.1 showing the maximum AOD over Germany, the model aerosol optical depth (AOD) indicates a period with higher SO4 concentrations starting with the update on 10 July, while all other components show no marked changes. Saharan dust events occurred from 9 - 16 June and from 24 June – 2 July. With respect to AOD, all aerosol components follow their usual seasonality.
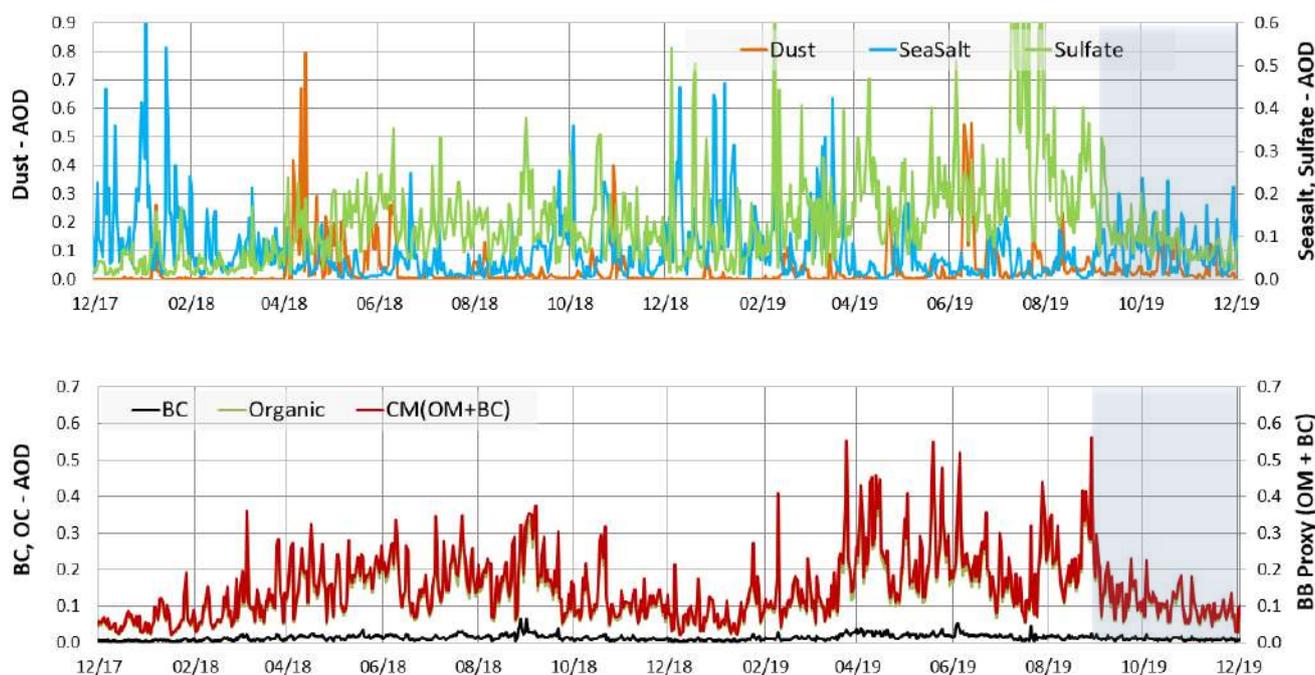
Figure 7.3.1: Maximum daily AOD over Germany for aerosols included in the IFS model from 12/2017 - 11/2019: sea salt (blue), dust (orange), sulphate (light green), black carbon (BC, black), organic matter (green), proxy for 'biomass burning' (as OC+BC - red). Note the different y-axes for the aerosol species.
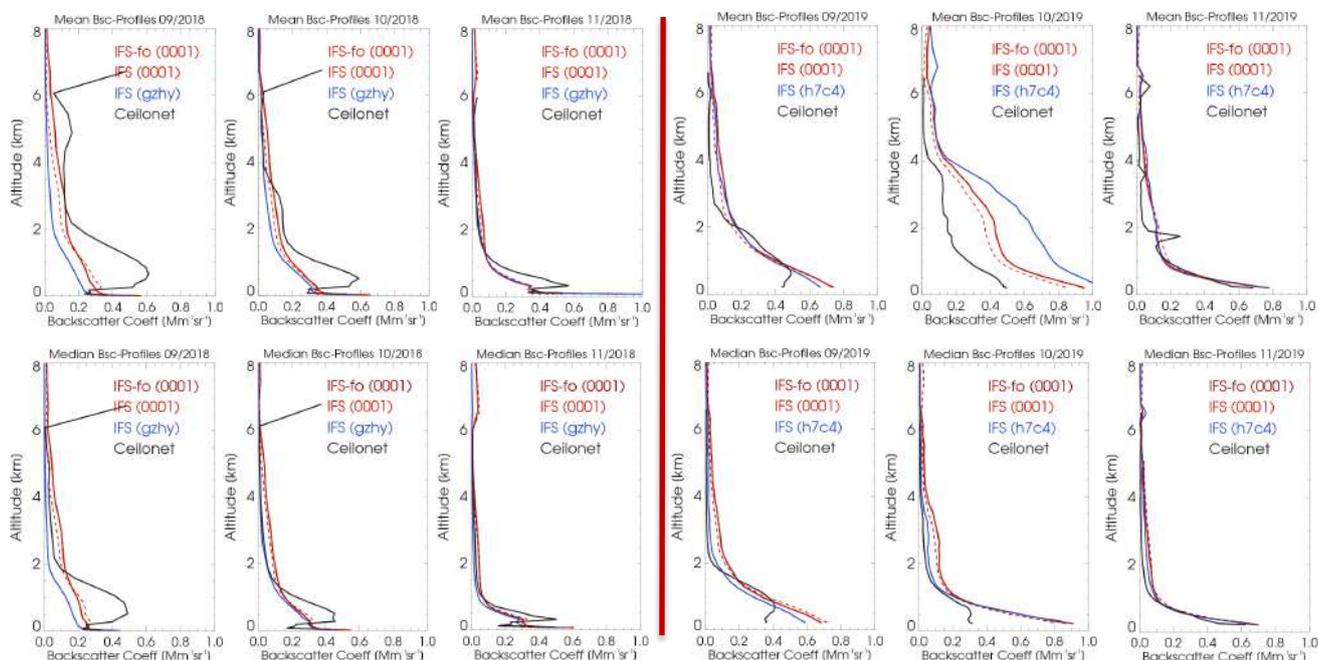
Figure 7.3.2: Left 2018, right 2019 (o-suite: red; control: blue). Monthly mean profiles (upper panel) and median profiles (lower panel) of backscatter coefficients from o-suite (red), control run (blue), and ceilometers (black) combined from 21 German stations in Sep 2019 to Nov 2019. The profiles are partly contaminated by remaining cloud artefacts.

Mean profiles:

In the new IFS cycle 46r1, implemented on 9 July 2019, nitrate NO_3 and ammonia NH_4 have been added and, likewise sulphate SO_4, have been coupled w.r.t homogeneous (gas-phase) and heterogeneous (particle phase) chemical processes. They contribute roughly 10-30% of aerosol mass in the rural central European PBL, as neutralized forms NO_3NH_4 or $(NH_4)_2SO_4$. Simultaneously, emissions of most aerosol components were significantly upgraded.

The negative bias of model BSC in the PBL has clearly decreased with transition from 60L to 137L model version. Further the 137L profiles (51L instead of 27L <8 km altitude) follow the step at the PBL top better than in the 60L version, except for Oct 19 where single events' overestimation causes larger deviations. O-suite and control are close to each other except again for Oct 19, where, however, the assimilation corrects the aerosol towards the observations and improves the profile as well. Both the model-observation correlation and the amplitude of the model vertical profile, coded in the standard deviation in the Taylor plots, have improved.
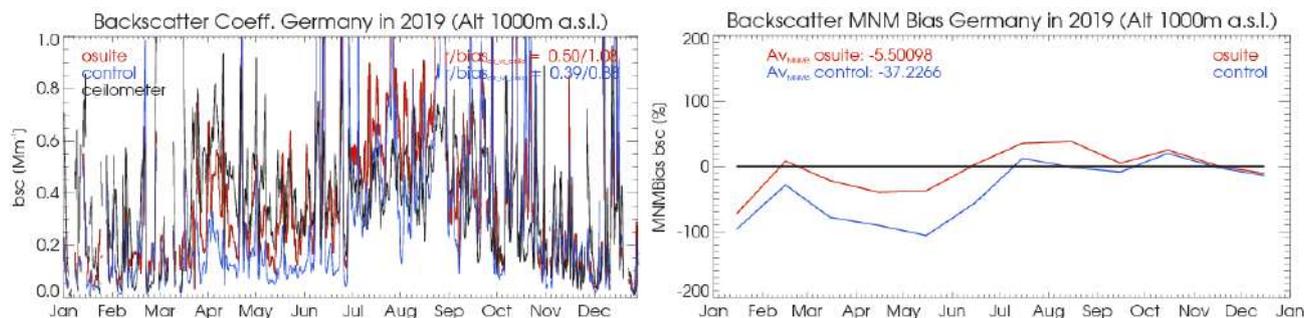
Figure 7.3.3. Time series of backscatter coefficients from o-suite (red), control run h7c4 (blue), and ceilometers (black) combined from 21 German stations in 2019 (left) and corresponding modified normalized mean bias (MNMB). Results refer to an altitude of 1000m above ground.
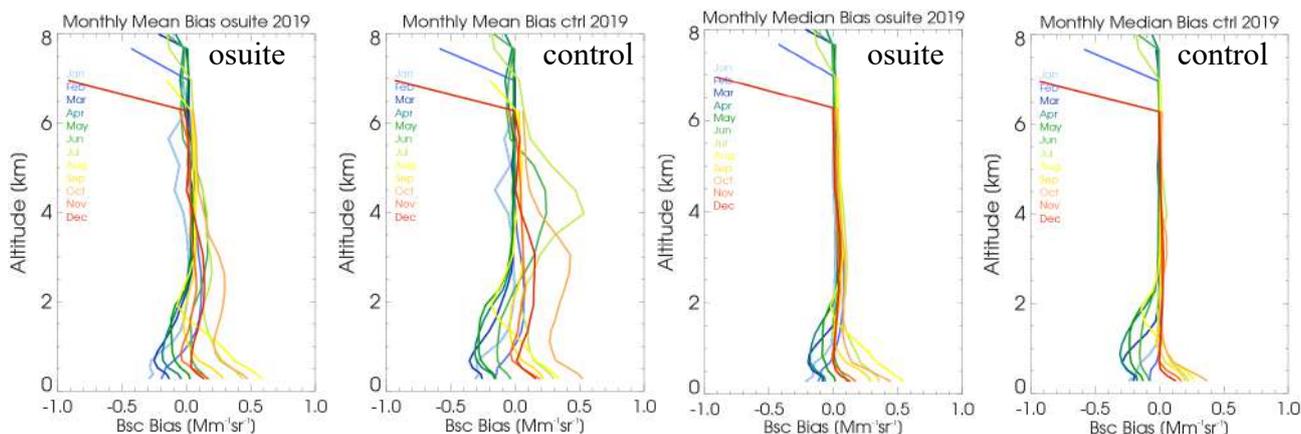
Figure 7.3.4 Monthly mean (left) and median (right) profiles of bias o-suite/ceilometer, control/ceilometer (run h7c4), combined from 21 German stations in 2019. In the upperpart, the profiles are partly contaminated by remaining cloud artefacts.

In Figure 7.3.4 the monthly mean bias profiles from the o-suite and control are summarized for the year 2019, as monthly mean (left two) and monthly median (right two) profiles, colour coded from blue for Jan to red for Dec. Systematic seasonal changes in the profile bias are most pronounced in the PBL where the aerosol is concentrated. Above 6 km cloud artefacts remaining in the ceilometer data hamper the comparison. Clearly, the biases are less for the o-suite than for the control runs. Comparison to the median profiles indicates that events determine most of the bias.

The high mean bias below 4 km in Oct results from Saharan dust events SDE (11 days with $AOD_{SD} >$ 0.05) which are over-estimated by the model. The vertical amplitude is thus larger in the mean profiles (more affected by events) than in the median profiles, which reflect more background conditions.
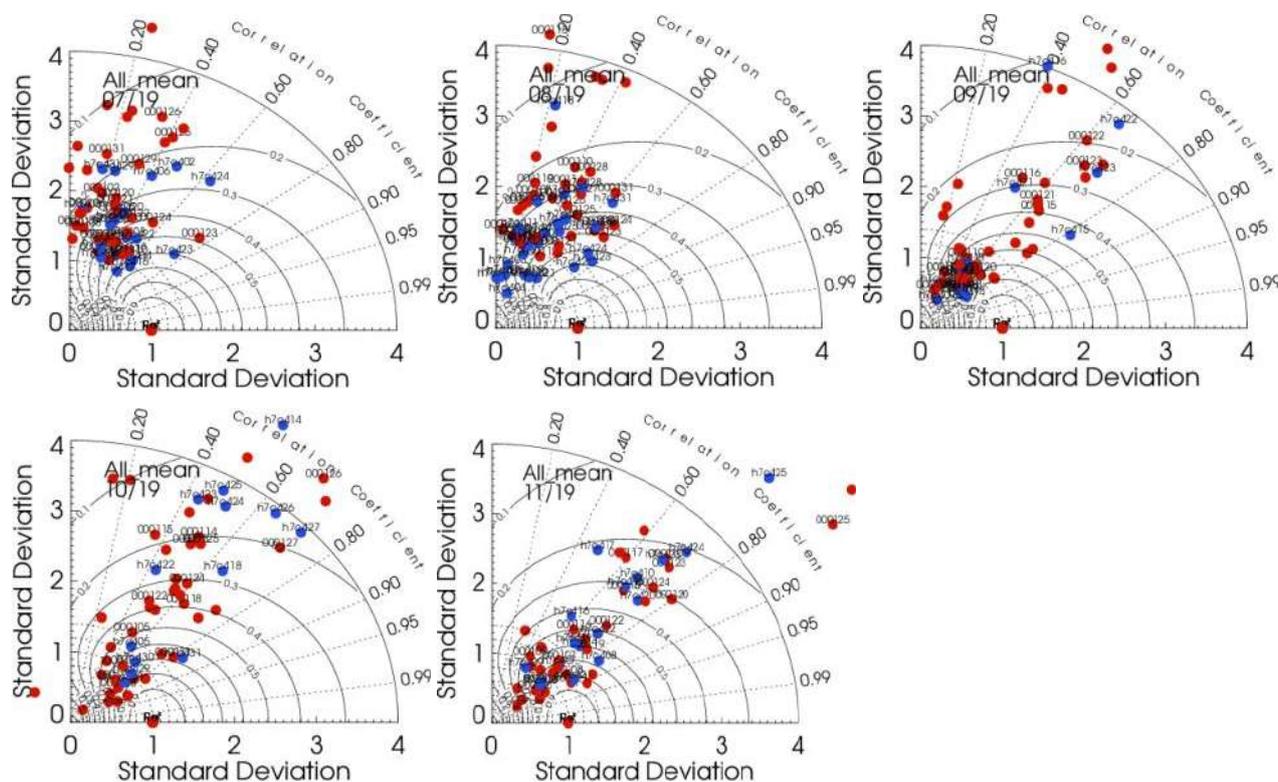
Fig. 7.3.5. Taylor polar plots with daily average standard deviation of vertical profiles vs correlation coefficient, averaged over 21 German ceilometer sites for Aug – Nov 2019. O-suite: red; control: blue.

It must be noted that the profiles 'IFS (0001 & h7c4)' are calculated preliminarily using mass-to-backscatter conversion coefficients from SO_4 for both NO_3 and NH_4, which is a rough approximation for the moment and will be replaced by correct values in future. The dashed profiles using the IFS-forward operator are, however, correct and confirm that this approximation does not change the main conclusions concerning the progress seen by the model upgrade in July 2019.

Taylor Plots:

The average coefficient of correlation between modelled and observed vertical backscatter profiles has improved from the 60L to the 137L model version and now ranges between r = 0.4 – 0.8 rather than r = 0.2-0.6 before. (Fig. 3.5.18). This is mainly due to better representation of the PBL top height. The absolute standard deviation (SD) are normalized to the SD of observations per day, as reference value at SD ≡ 1.0. In autumn 2019, o-suite (red dots) and control run (blue dots) show similar correlation with the observations, but the standard deviation (i.e. the vertical amplitude) is captured better by the o-suite. Fig. 3.5.17 shows, that the Saharan dust events are quantified better in the o-suite than in control. There is a large day-to-day and also a seasonal variation of the performance. It must be noted, that small vertical displacements decrease the correlation coefficient, although the SD plumes are mostly reproduced at geographically truthful positions.

Summary

Conclusion concerning changes due to the model upgrade in July 2019 are:

- The negative bias of model BSC in the PBL has reduced.

- The bias of model BSC during events in the FT has not changed notably.

- The step at the top of the PBL is captured notably better with 137 levels than with 60 levels (51L instead of 27L <8 km altitude), same for o-suite and control.

- The absolute values of the model vertical profile, dominated by the PBL, are now closer to the observations.

Daily averages of Pearsons correlation coefficients cluster around r= 0.4-0.8 for the 137L model, in contrast to r=0.2-0.6 for the 60L version before.

7.4 Aerosol validation over Europe and the Mediterranean

Three-hourly aerosol optical depth (AOD) and surface concentration (PM10 and PM2.5) from the o-suite and control run have been validated for the period 1 September 2019 – 30 November 2019 against AERONET direct-sun cloud-screened observations.

Aerosol optical depth over the Mediterranean

The CAMS o-suite does reproduce the daily variability of AERONET direct-sun observations. In Western, Central and Eastern Mediterranean, the correlation coefficient increases from [0.70, 0.64 and 0.64], to [0.78, 0.71 and 0.68], respectively for control and o-suite during autumn (see Figure 7.4.1). In general, both CAMS experiments overestimated the AERONET observations in the Mediterranean basin in control (MB of 0, 0.01 and -0.01 for Western, Central and Eastern Mediterranean regions respectively) and o-suite (MB of 0.02, 0.03 and 0.01 for Western, Central and Eastern Mediterranean regions respectively) as shown in Figure 7.4.1. The highest peaks on CAMS AOD simulations are linked to desert dust intrusions occurring during the whole season in the whole Mediterranean basin. This is shown in the Barcelona (Spain), Rome (Italy) and Sede Boker (Crete) AERONET sites (see Figure 7.4.2). In Barcelona AOD values up to 1.0 are observed. In the second half of October, elevated AOD values are observed in Sede Boker (up to 0.5) that are not associated with natural aerosols (i.e. dust or sea-salt). This can be associated to the latest model upgrade with the improvements implemented for the sulphur cycle.
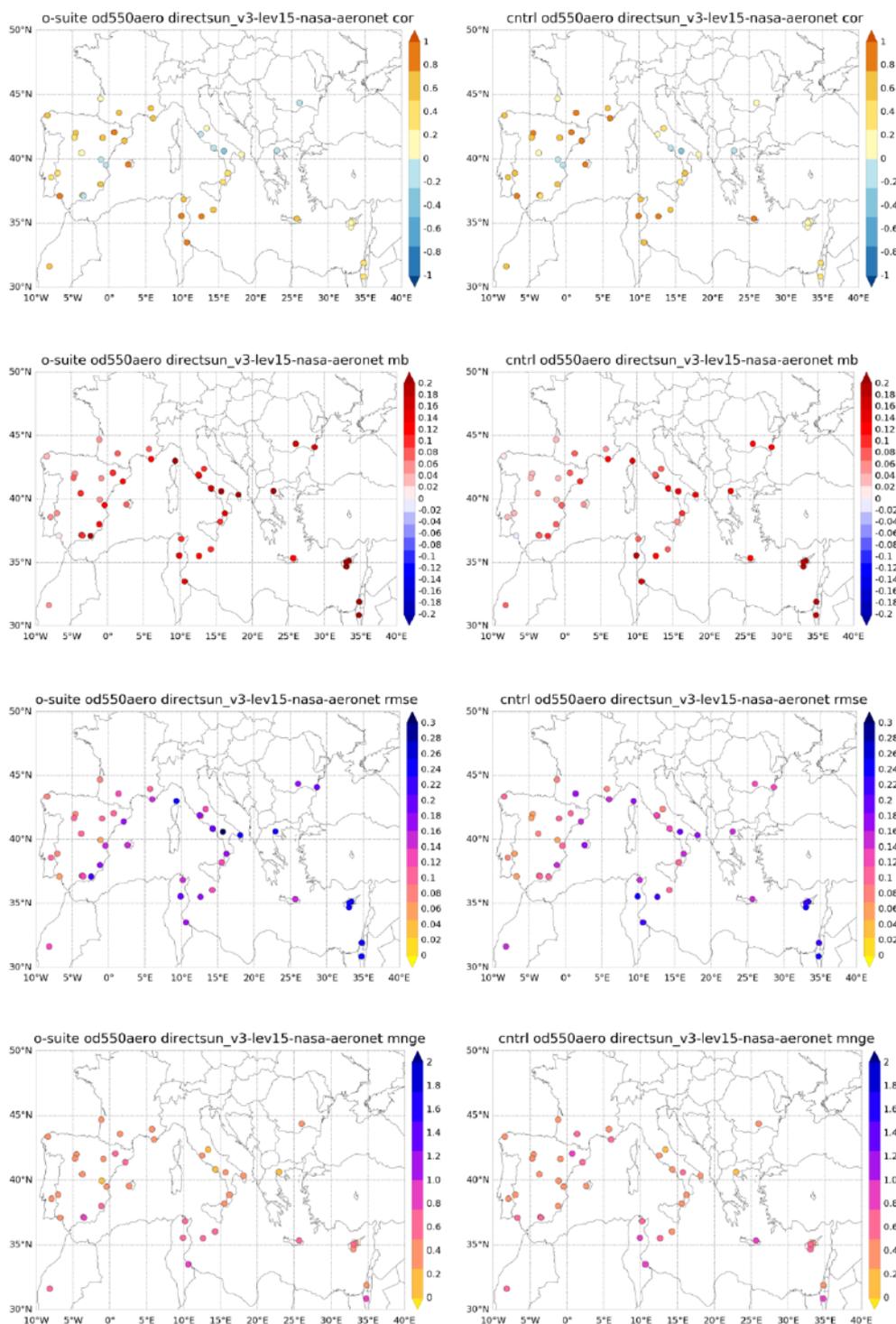
Figure 7.4.1: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite and control for the study period. AOD from AERONET direct-sun is the reference.
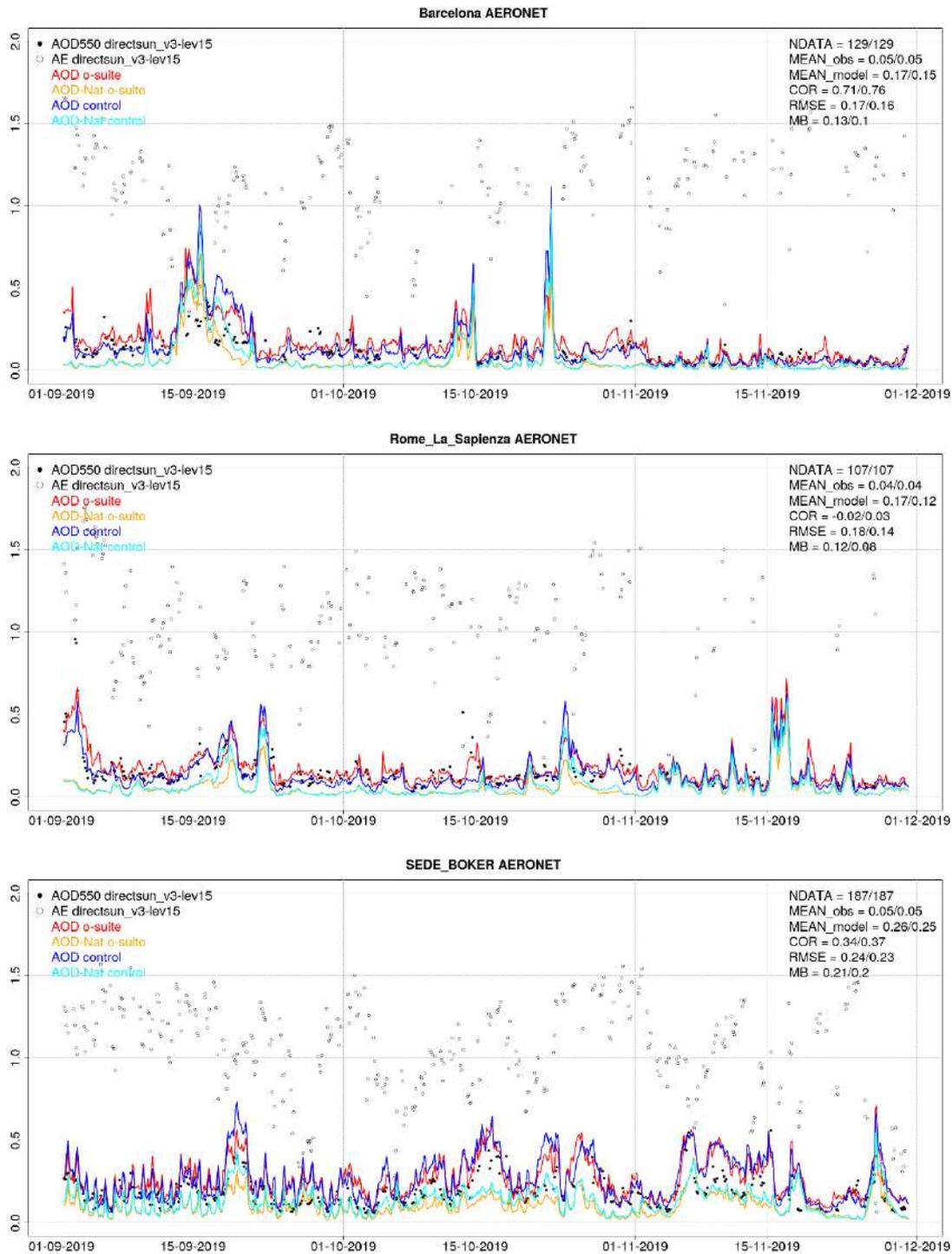
Figure 7.4.2: AOD from AERONET (black dot), AOD o-suite (red line), AOD control (blue line), AOD-Nat o-suite (orange line), AOD-Nat control (cyan line), for the study period over Barcelona (Spain), Rome La Sapienza (Italy) and SEDE BOKER (Crete). AOD-Nat corresponds to the natural aerosol optical depth that includes dust and sea-salt. Skill scores per each individual site and model (o—suite/control) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN model, COR, RMSE, MB).
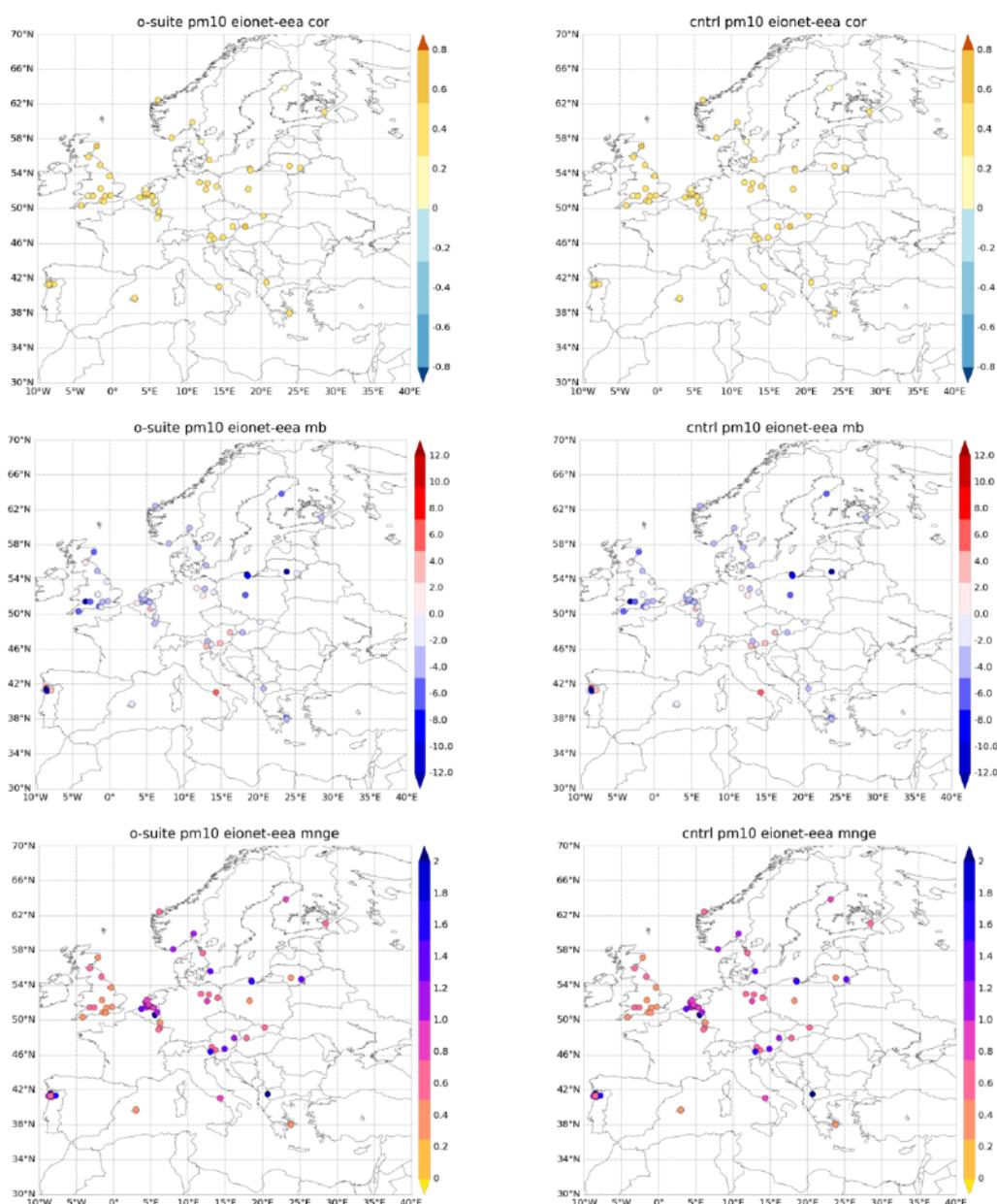
Figure 7.4.3: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite and control for the study period. PM10 from EIONET are the reference. Only background suburban and rural available stations are displayed.

Surface aerosol concentrations in Europe

For autumn, CAMS o-suite and control show similar skill scores for PM10 and PM2.5 in comparison with EIONET-Airbase observations (see Figure 7.4.3). The CAMS products tend to underestimate the PM10 and PM2.5 EIONET-Airbase observations (see Figure 7.4.3 and Figure 7.4.4). The o-suite shows higher overestimations in PM10 and PM2.5 (with MB of -2.93 and -1.64 µg/m³, respectively) than control (with MB of -4.43 and -2.68 µg/m³, respectively). Overestimations are observed in Southern European sites, particularly in Italy for PM10 (see Figure 7.2.5).
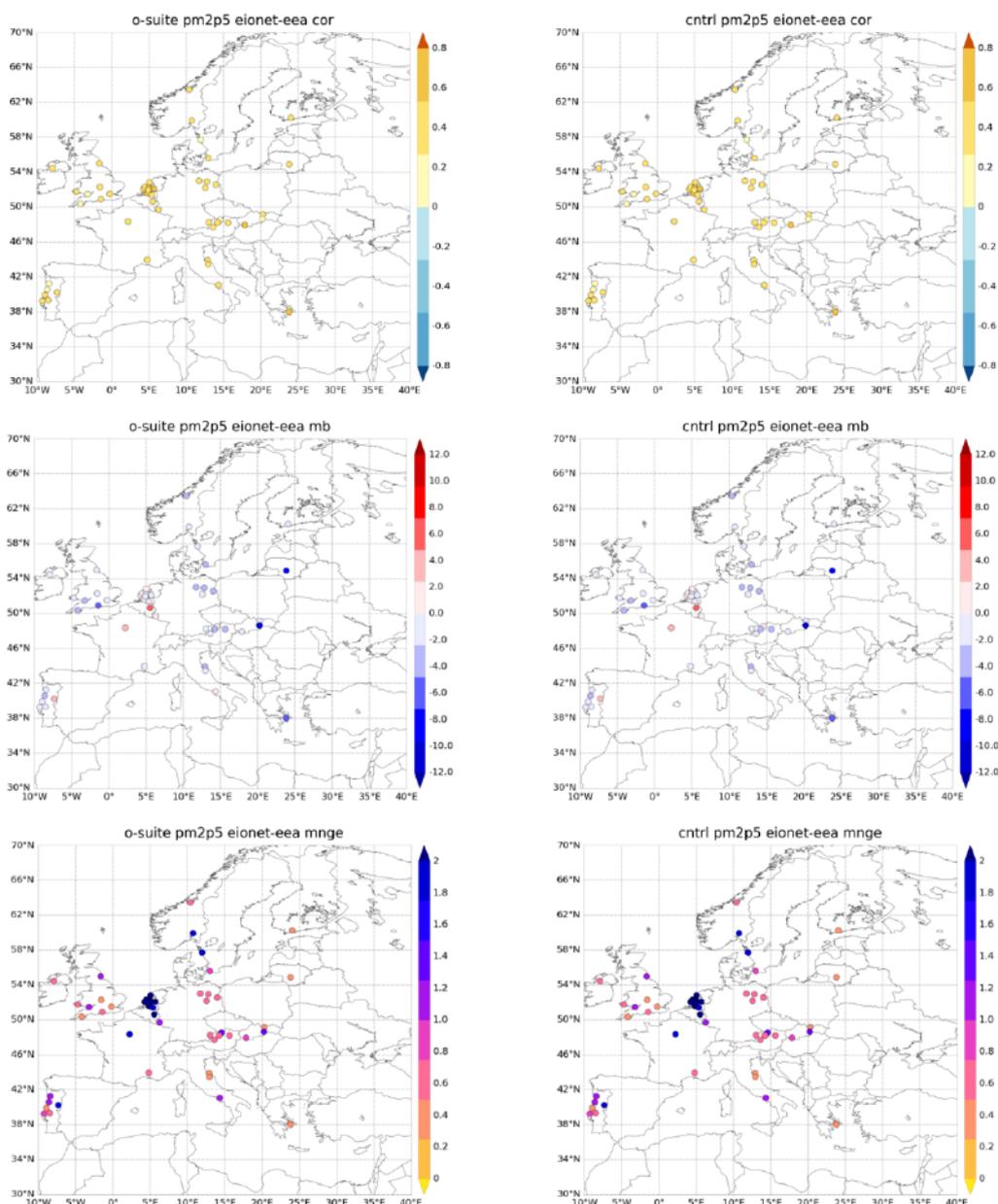
Figure 7.4.4: Skill scores (correlation coefficient, MB, RMSE and FGE) for 24-hour forecasts of CAMS o-suite and control for the study period. PM2.5 from EIONET are the reference. Only background suburban and rural available stations are displayed.

The upgrade of the CAMS o-suite during July 2019 led to an increase of the coarse particles at the surface, as shown in Figure 7.2.5. During the whole study period peaks above 50 μg/m³ for PM10 are observed.

ES0296A(3.01488ºE; 39.74731ºN) AIRBASE

NDATA = 627/627 || NDATA = NA/NA
MEAN_obs = 18.15/18.15 || MEAN_obs = NA/NA
MEAN_model = 17.76/17.82 || MEAN_model = NA/NA
COR = 0.44/0.43 || COR = NA/NA
RMSE = 11.52/13.97 || RMSE = NA/NA
MB = -0.39/-0.33 || MB = NA/NA

IT1166A(13.01111ºE; 46.40194ºN) AIRBASE

NDATA = 488/488 || NDATA = NA/NA
MEAN_obs = 11.93/11.93 || MEAN_obs = NA/NA
MEAN_model = 14.39/13.02 || MEAN_model = NA/NA
COR = 0.23/0.25 || COR = NA/NA
RMSE = 17.02/16.81 || RMSE = NA/NA
MB = 2.46/1.09 || MB = NA/NA

PT01046(-8.29639ºE; 41.44972ºN) AIRBASE

NDATA = 618/618 || NDATA = NA/NA
MEAN_obs = 14.77/14.77 || MEAN_obs = NA/NA
MEAN_model = 11.24/10.7 || MEAN_model = NA/NA
COR = 0.03/0.02 || COR = NA/NA
RMSE = 15.15/15.19 || RMSE = NA/NA
MB = -3.53/-4.07 || MB = NA/NA

Figure 7.4.5: PM10 and PM2.5 Airbase observations (black and grey dots, respectively), PM10 and PM2.5 o-suite (red and orange lines, respectively) and PM10 and PM2.5 control (blue and cyan lines, respectively) for the study period over ES1691A (Spain), IT1166A (Italy) and PT01046 (Portugal).

8. Stratosphere

8.1 Validation against ozone sondes

In this section, we present the results of the stratospheric ozone evaluation against ozone soundings from the NDACC, WOUDC, NILU and SHADOZ databases. The sondes have a precision of 3-5% (~10% in the troposphere for Brewer Mast) and an uncertainty of 5-10%. For further details see Cammas et al. (2009), Deshler et al. (2008) and Smit et al (2007). Model profiles of the o-suite are compared to balloon sondes measurement data of 44 stations for the period January 2013 to November 2019 (please note that towards the end of the validation period fewer soundings are available). As C-IFS-CB05 stratospheric composition products beyond O_3 in the o-suite is not useful we provide only a very limited evaluation of the control experiment. A description of the applied methodologies and a map with the sounding stations can be found in Eskes et al. (2019). Please note that recent scientific findings (https://tropo.gsfc.nasa.gov/shadoz/Archive.html, Thompson et al., 2017; Witte et al., 2017; 2018, Stauffer, et al. in preparation 2020) show a drop-off in Total Ozone at various global ozone stations in comparison with satellite instruments. This drop-off amounts between 5-10% for stratospheric ozone. Changes in the ECC ozone instrument are associated with the drop-off, but no single factor has been identified as cause yet.

The o-suite shows MNMBs within the range ±12%, for all regions and months (some exceptions with MNMBs of up to ±18% for single months in the high latitude regions). Figure 8.1.1. shows the results for the past year.

Fig. 8.1.2 compares the averaged profiles in each region during October 2019. The vertical distribution of stratospheric ozone is quite well represented for all regions by the o-suite, with little overestimation in all latitude bands (MNMBs between -1 to +14% for SON 2019). The control run shows a strong overestimation of stratospheric ozone in the stratosphere and UTLS.
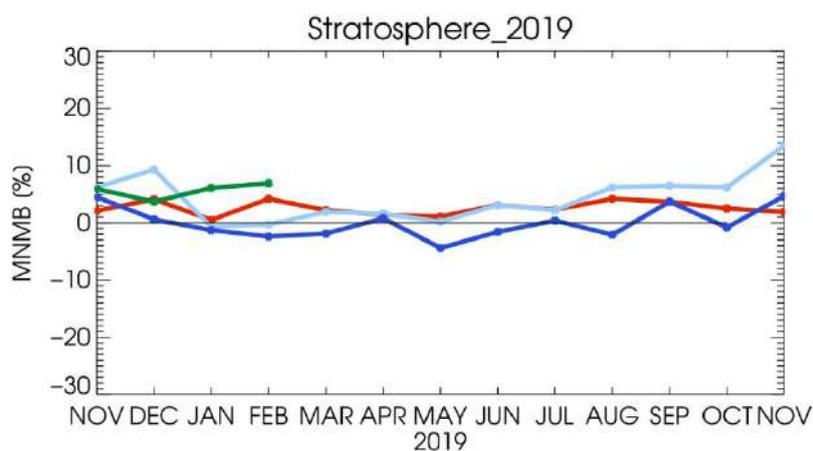
Figure 8.1.1: MNMBs (%) of ozone in the stratosphere from the o-suite against aggregated sonde data in the Arctic (light blue), Antarctic (dark blue) northern midlatitudes (red) and tropics (green). Period November 2018 to November 2019. The stratosphere is defined as the altitude region between 60 and 10 hPa in the tropics and between 90 and 10 hPa elsewhere.
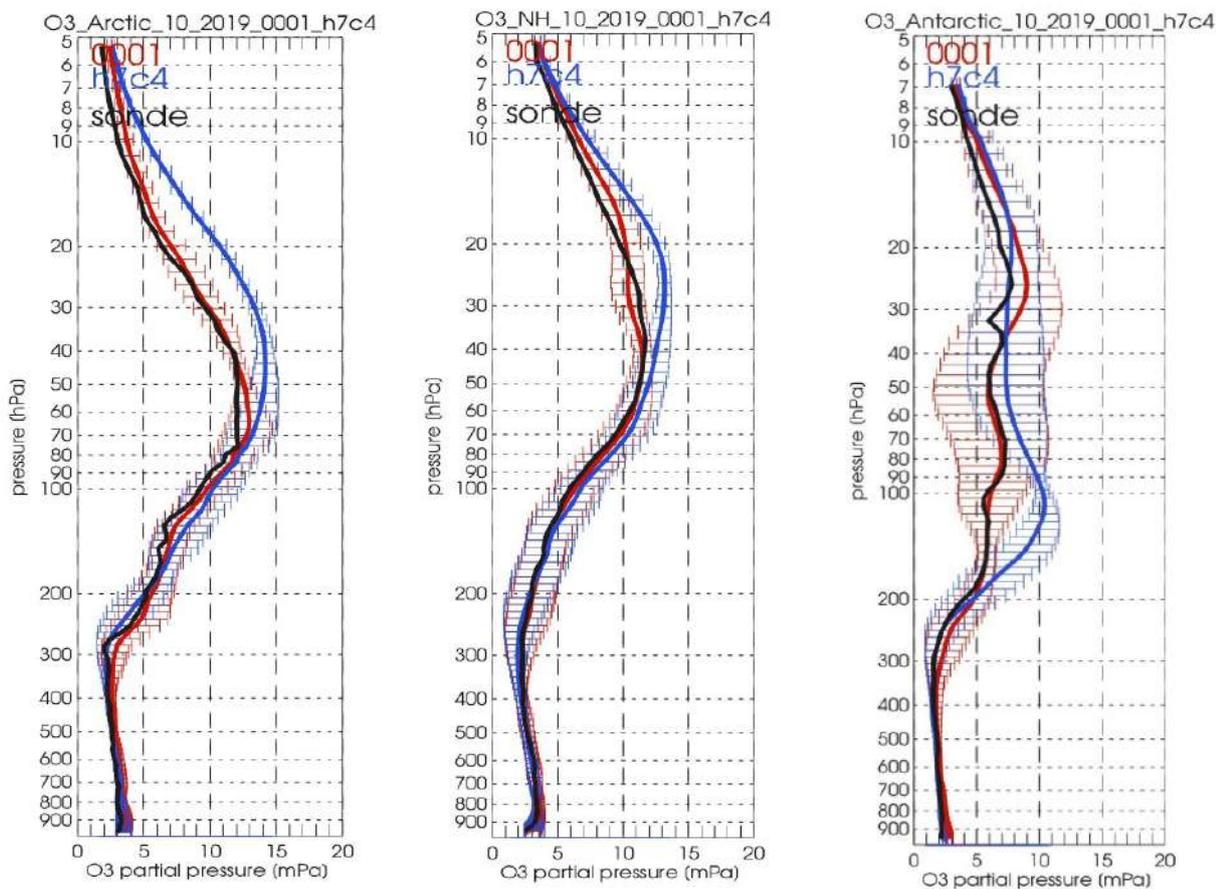
Figure 8.1.2: Comparison between mean O₃ profiles (units: mPa) of o-suite (red), and control (blue) in comparison with observed O₃ sonde profiles (black) for October 2019 for the various latitude bands: Arctic, NH-mid latitudes and Antarctic.

8.2 Validation against observations from the NDACC network

UVVIS and FTIR stratospheric columns

Since the start of the CAMS27 project, the number of UVVIS Zenith ozone measurements have increased on NDACC. Currently sixteen sites provided data in the recent quarter allowing for a representative picture on the latitude dependence of the model data.

The systematic uncertainty of the UVVIS measurements is typically 5%, hence the relative biases for most sites for both the AN and 1d FC of the o-suite are very close to each other and within the uncertainty ranges, see Figure 8.2.1. The averaged bias for the 16 UVVIS sites is comparable to the measurement uncertainty of 5%, the averaged correlation is above 0.9.

The correlations between the sites and the model are presented in the Taylor diagrams in Figure 8.2.2. Again, the o-suite AN and 1d FC perform very similarly in correlation coefficients.

Figure 8.2.3 depicts the FTIR stratospheric columns showing a discontinuity in the o-suite 1d FC model for the tropical sites (Mauna Loa, Altzomoni and Reunion) in the June 2016 model update. The worse performance of the tropical sites is also seen in lower correlations in Figure 8.2.2)
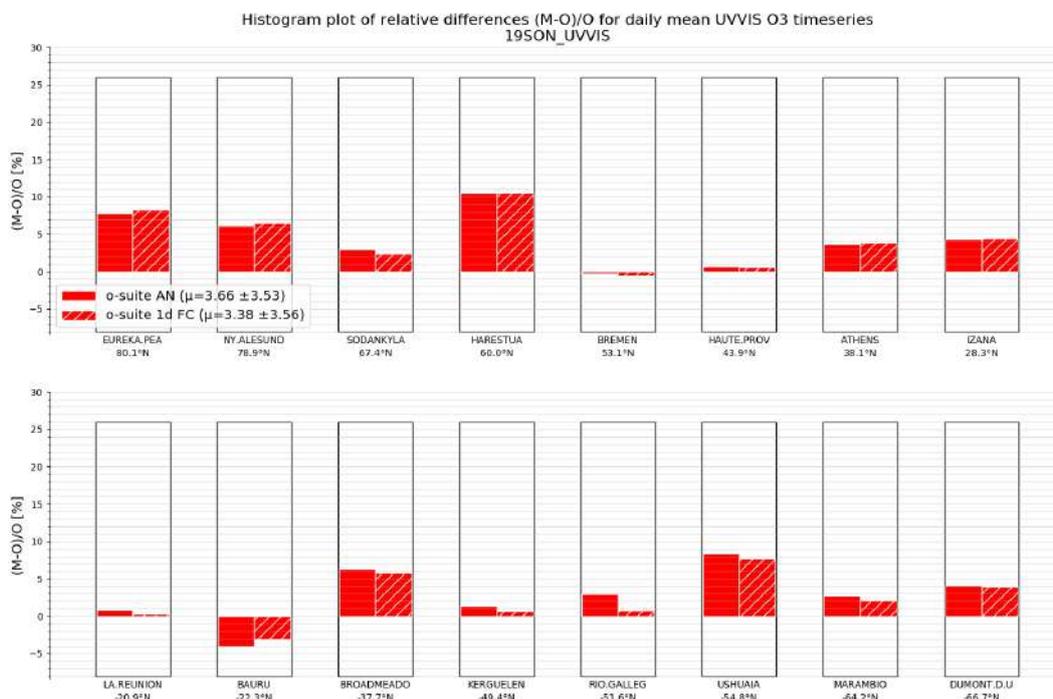
Figure 8.2.1 Relative biases during quarter SON 2019 for 16 UVVIS stations measuring stratospheric ozone columns with ZENITH measurement geometry (stations sorted with decreasing latitude). The overall relative bias is positive for all latitudes and comparable to the typical measurement uncertainty of 5% for most sites.

Profile comparison using LIDAR and MWR

In this section we present a comparison between the CAMS o-suite and control products against MWR and LIDAR observations from the NDACC network. A detailed description of the instruments and applied methodologies for all NDACC instruments can be found at http://nors.aeronomie.be. MWR (microwave) at Ny Alesund (79°N, 12°E, Arctic station) and Bern (47°N, 7°E, northern midlatitude station). LIDAR at Observatoire Haute Provence (OHP), France (43°N, 5.7°E, altitude 650m), Hohenpeissenberg, Germany (47°N, 11°E, altitude 1km) and Mauna Loa, Hawaii (19.5°N, 204°E, altitude 3.4km).

At OHP, Hohenpeissenberg and Mauna Loa (LIDAR, see Figure 8.2.4), the o-suite slightly overestimates the observed ozone (<10%) between 25km and 35km. The uncertainty on the LIDAR concentration increases with altitude and above 35km the observed differences are comparable to the measurement uncertainty (>10%, see http://nors.aeronomie.be/projectdir/PDF/NORS_D4.2_DUG.pdf).
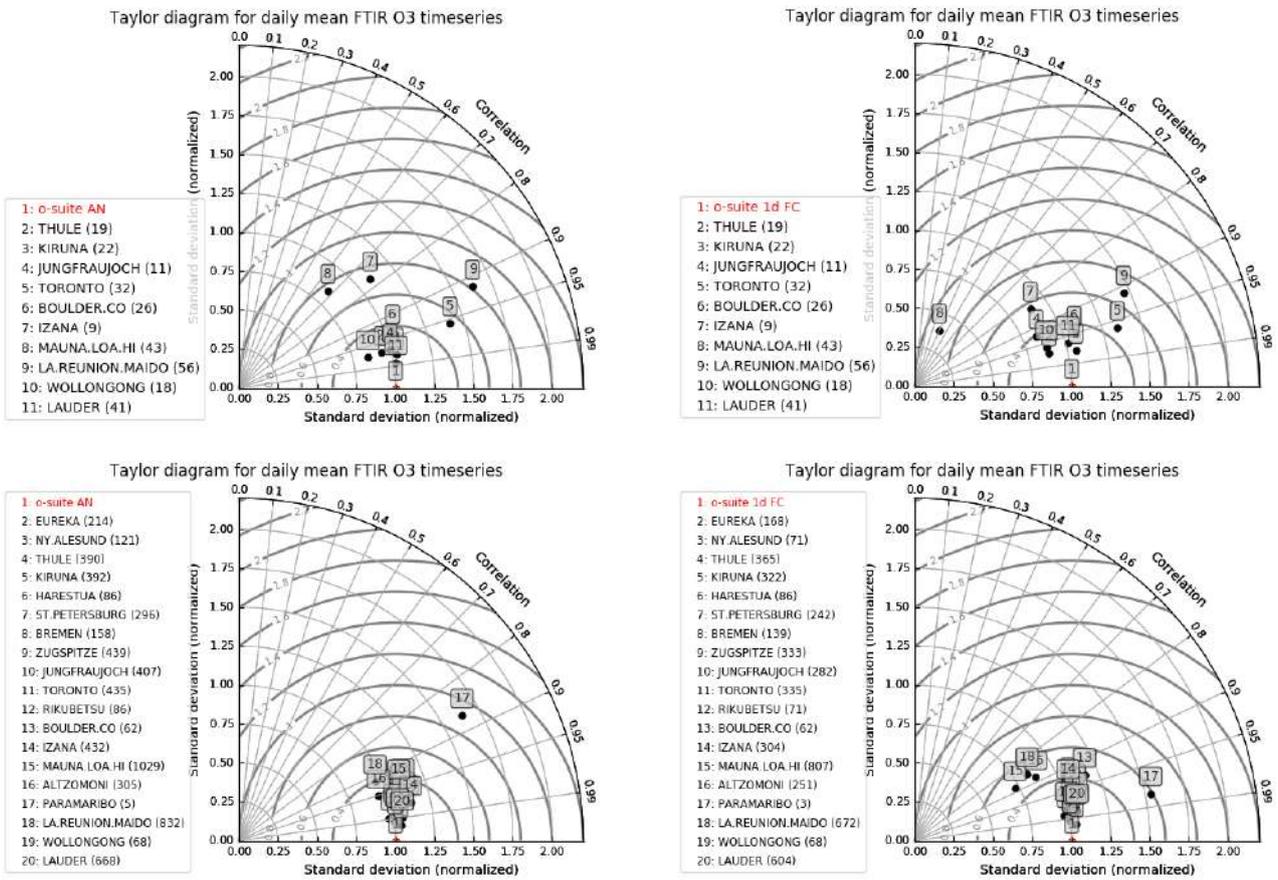
Figure 8.2.2 Taylor diagrams relating the standard deviations for the model and GB stratospheric column time series and their correlation for the time period SON 2019. All timeseries are normalized such that the std of the model is 1. The performance for the o-suite AN is slightly better (averaged correlation is 0.91 for FTIR and 0.85 for UVVIS) compared to the 1d FC (averaged correlation is 0.88 for FTIR and 0.84 for UVVIS). Again, the correlation for the tropical sites are worse in the 1dFC compared to the AN.
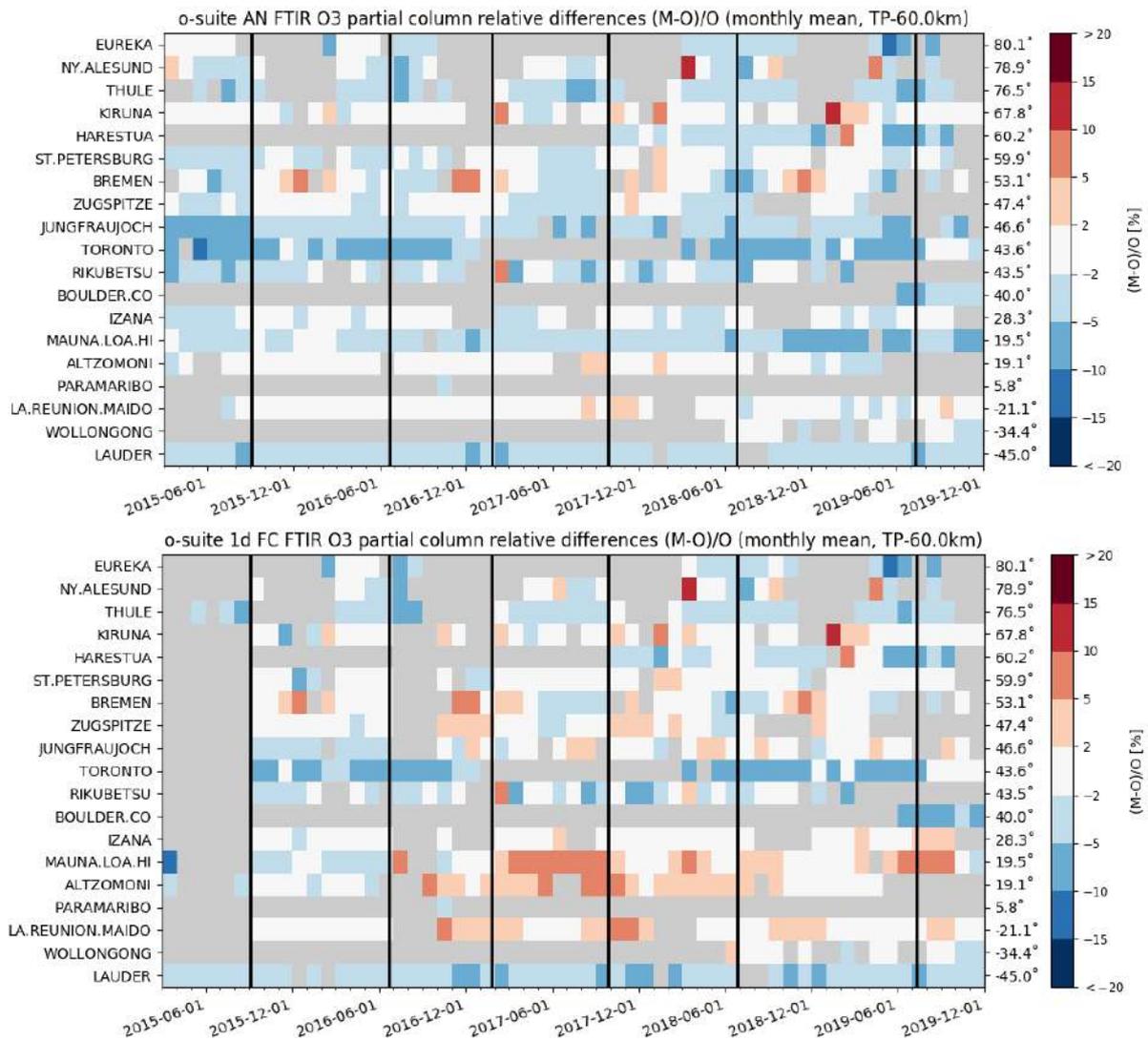
Figure 8.2.3 Time series of monthly mean relative differences for stratospheric FTIR columns along with model cycle updates (black vertical lines) (o-suite AN top, o-suite 1d FC bottom). The stratospheric FTIR columns for the tropical sites at Izana, Mauna Loa, Altzomoni and Reunion show a higher overestimation for the 1dFC compared to the AN.

MWR.O3 VMR profile differences (M-O)/O
(weekly mean, 25.0km - 65.0km, o-suite AN, NY.ALESUND (lat.=78.9°), 2016-02-01 till 2019-10-12, 9128 meas.)

MWR.O3 VMR profile differences (M-O)/O
(weekly mean, 25.0km - 65.0km, o-suite AN, BERN (lat.=47.0°), 2015-03-01 till 2019-11-30, 12788 meas.)

LIDAR.O3 number density profile differences (M-O)/O
(weekly mean, 20.0km - 45.0km, o-suite AN, HOHENPEISSENBERG (lat.=47.8°), 2015-03-06 till 2019-11-30, 532 meas.)

Figure 8.2.4: Comparison of the weekly mean profile bias between the O_3 mixing ratios of o-suite and the NDACC station at Ny Alesund, Bern, Hohenpeissenberg, OHP and Mauna Loa. For the LIDAR stations, the measurement uncertainty above 35km is comparable to the observed profile bias.

8.3 Comparison with dedicated systems and with observations by limb-scanning satellites

This section compares the output of the o-suite for the last period with observations by limb-scanning satellite instruments, using the methodology described by Lefever et al. (2015). We also include the comparisons for the o-suite 4th day forecasts (96h to 120h) of stratospheric ozone. These forecasts are represented by dotted lines in the figures.

All datasets are averaged over all longitudes and over the three most interesting latitude bands for stratospheric ozone: Antarctic (90°S-60°S), Tropics (30°S-30°N) and Arctic (60°N-90°N). In order to provide global coverage, the two mid-latitude bands (60°S-90°S and 60°N-90°N) are also included in some comparisons with satellite observations.

o3 relative bias against observations: 10-30hPa mean from 20161201 to 20191201
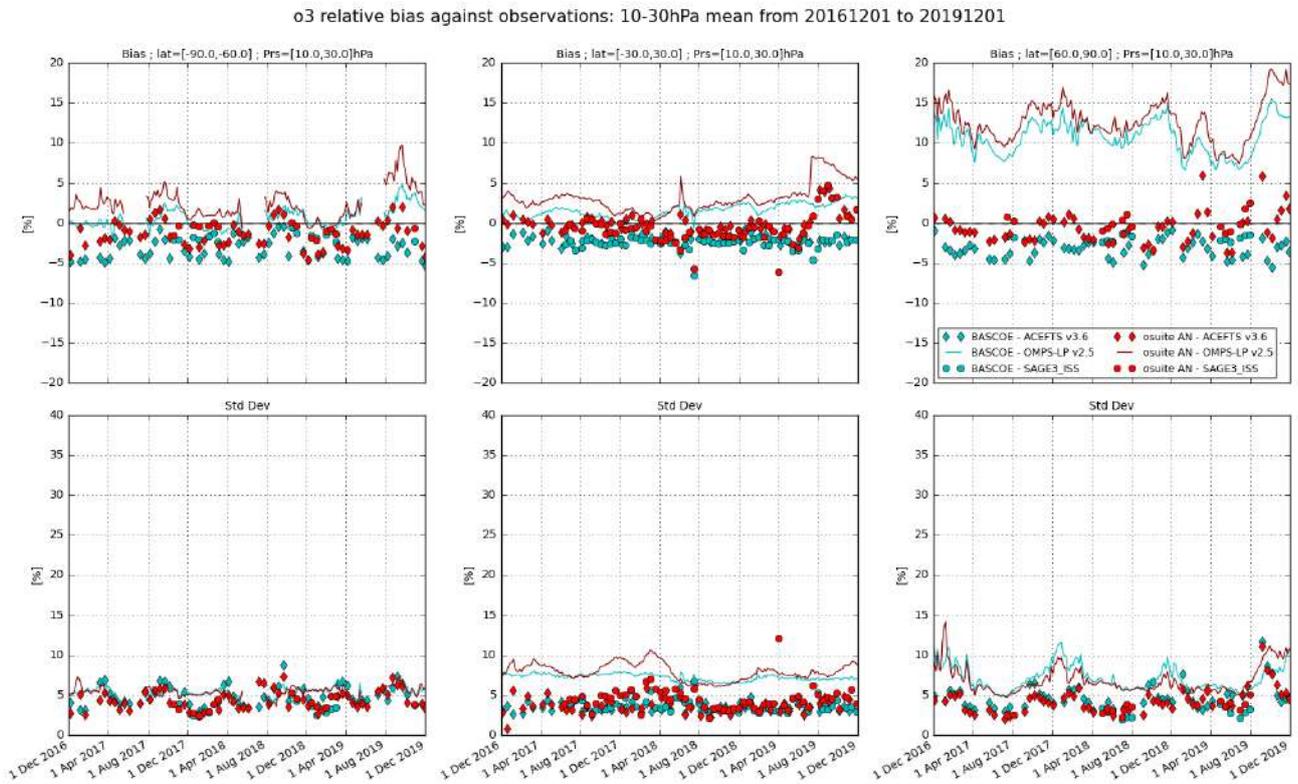
Figure 8.3.1: Time series comparing models to observations for the period 2016-12-01 to 2019-12-01 in the upper stratosphere (10-30hPa averages): o-suite analyses (red) and BASCOE (cyan) vs OMPS-LP (solid), ACE-FTS (diamonds) and SAGE-III (bullets). Top row: normalized mean bias (model-observation)/observation (%); bottom row: standard deviation of relative differences (%).

The level-2 data from limb scanning instrument used in this section are:

- ACE-FTS version 3.6, on board SCISAT-1

- SAGE-III version 5.1, on board the International Space Station (ISS); among the 3 different ozone profiles delivered by the solar occultation (denoted Mesospheric, MLR and AO3), we use the AO3 retrieval which is recommended by the mission science team.

- OMPS-LP version 2.5, on board NPP

For reference, we include also the BASCOE analyses which are very constrained by the AURA MLS offline profiles.

Figure 8.3.1 to 8.3.3 present, in the upper row, the timeseries over the last 36 months of the bias of the o-suite against the three satellite measurements for respectively three layers of the stratosphere (10-30hPa upper, 30-70hPa middle, and 70-100hPa lower and UTLS); the bottom row of the figures shows the standard deviation of the differences and can be used to evaluate the random error in the analyses.
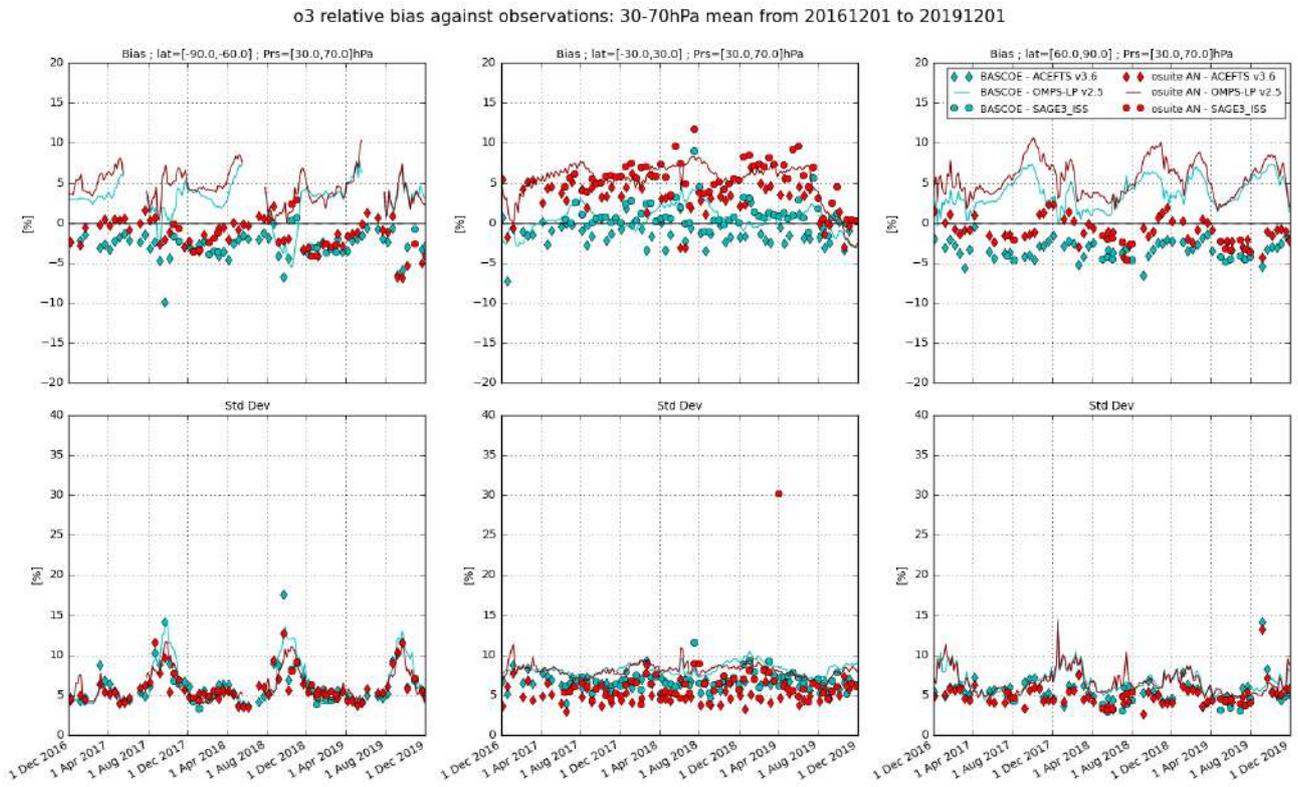
Figure 8.3.2: Time series comparing models to observations for the period 2016-12-01 to 2019-12-01 in the middle stratosphere (30-70hPa averages): o-suite analyses (red) and BASCOE (cyan) vs OMPS-LP (solid), ACE-FTS (diamonds) and SAGE-III (bullets). Top row: normalized mean bias (model-observation)/observation (%); bottom row: standard deviation of relative differences (%).

In the tropics for the 70-100hPa region, the comparison with all instruments is unreliable (highly scattered bias and large standard deviations)

The agreement with ACE-FTS is good: the bias is generally within ±5% for all regions.

The SAGE-III onboard ISS provide observations since June 2017. The latitudinal coverage is more limited than ACE-FTS; the polar regions are not covered for long periods of time (data available only in the summer). Where available, the agreement of the o-suite with SAGE-III is good, with biases similar to those observed against ACE-FTS, except in the tropics in the 30-70hPa region where they are more positive (3-13%).

Compared to OMPS-LP, there is an almost systematic overestimation by the o-suite; the biases are more variable and more marked than for the other instruments (10% to 15% in the north polar at 10-30hPa region, up to 10% at 30-70hPa and up to 20% at 70-100hPa).
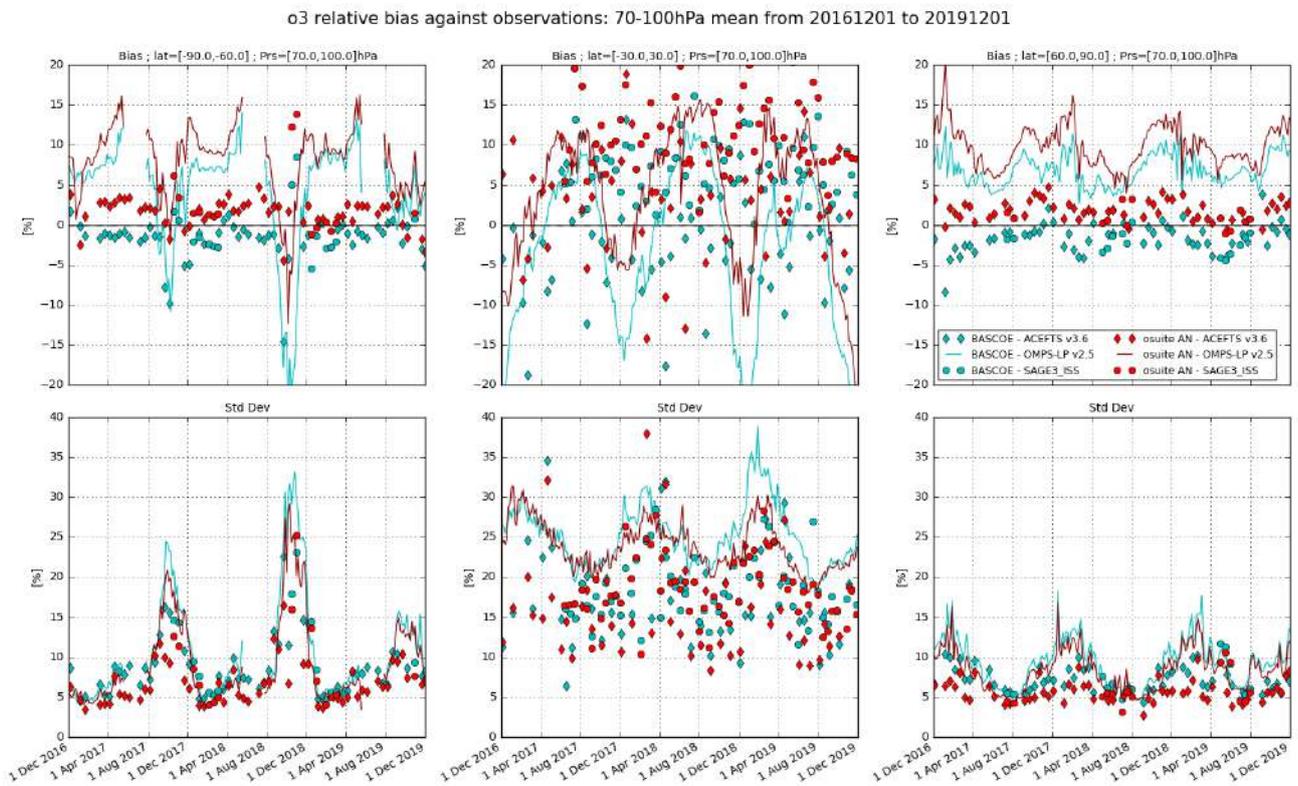
Figure 8.3.3: Time series comparing models to observations for the period 2016-12-01 to 2019-12-01 in the lower stratosphere (70-100hPa averages): o-suite analyses (red) and BASCOE (cyan) vs OMPS-LP (solid), ACE-FTS (diamonds) and SAGE-III (bullets). Top row: normalized mean bias (model-observation)/observation (%); bottom row: standard deviation of relative differences (%).

The bias of BASCOE against the satellite observations for the considered regions is systematically lower but follows a similar evolution as the o-suite.

Figure 8.3.4 to 8.3.7 display vertical profiles of the relative biases between the o-suite or BASCOE and the satellite measurements. The difference is averaged over the most recent 3-month period considered in this validation report, i.e. September to November 2019.

The bias against each instrument remains within ±10% between 20km and 35km.

All o-suite profiles present a common feature of a slight overestimation at around 30km, followed by a stronger underestimation at around 40km, which is evidenced in the 4th day forecast.

At the higher part of the south polar profiles, an overestimation of ~1.5 ppm appears above 55km compared to ACE-FTS and 0.5 hPa compared to MLS.

It must be noted that the different instruments have a variety of spatial and temporal coverage: for a 3 month period and over the latitude bands considered, OMPS and Aura MLS provide daily data with more than 40000 valid profiles, while ACE-FTS provides around 700 profiles in the polar region and 200 profiles in the tropics, and SAGE-III around 800 profiles in each mid-latitude band and the tropics, and only around 100 profiles in the south polar region (and only close to 60°).
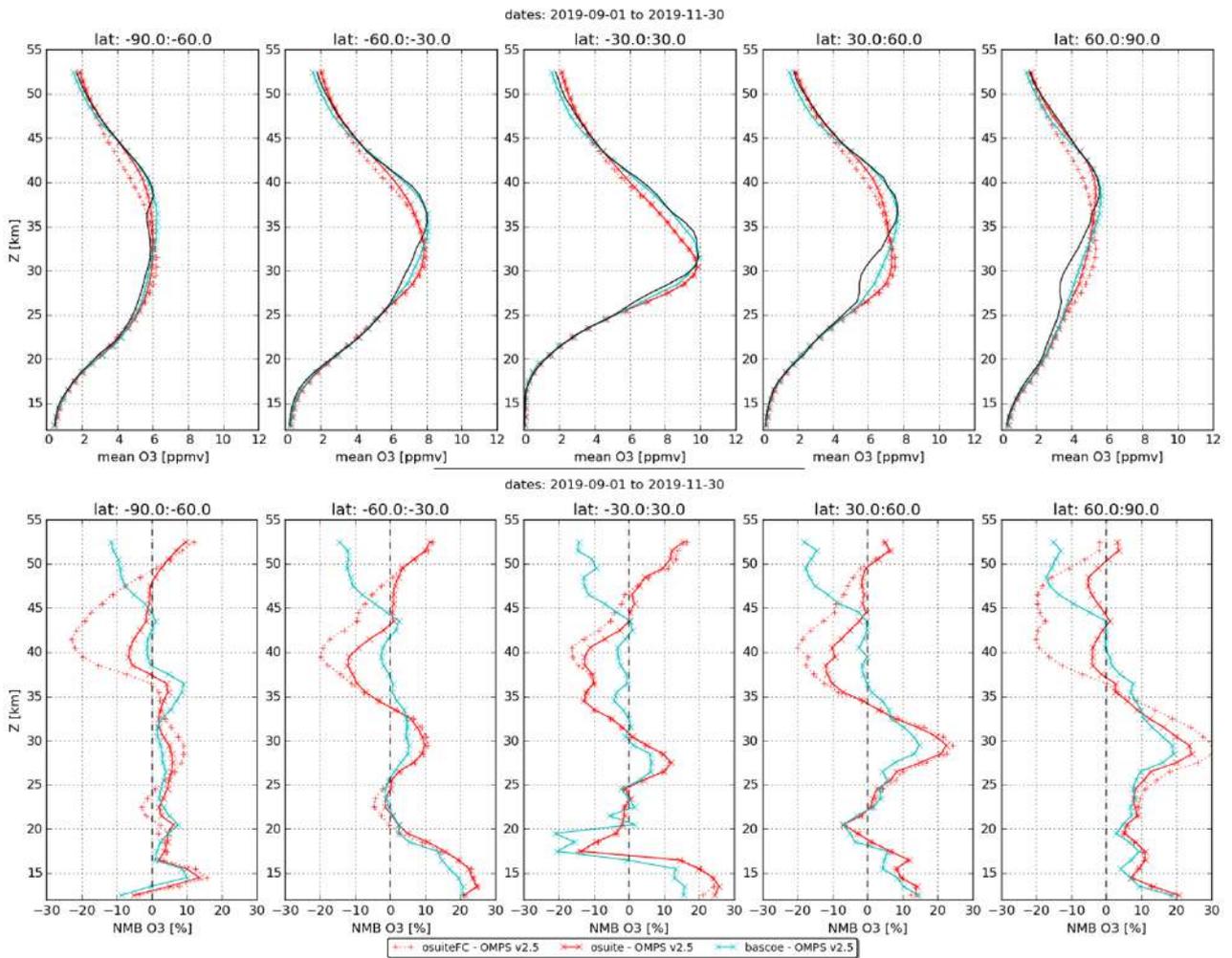
Figure 8.3.4: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with OMPS-LP v2.5 observations for the period SON 2019.
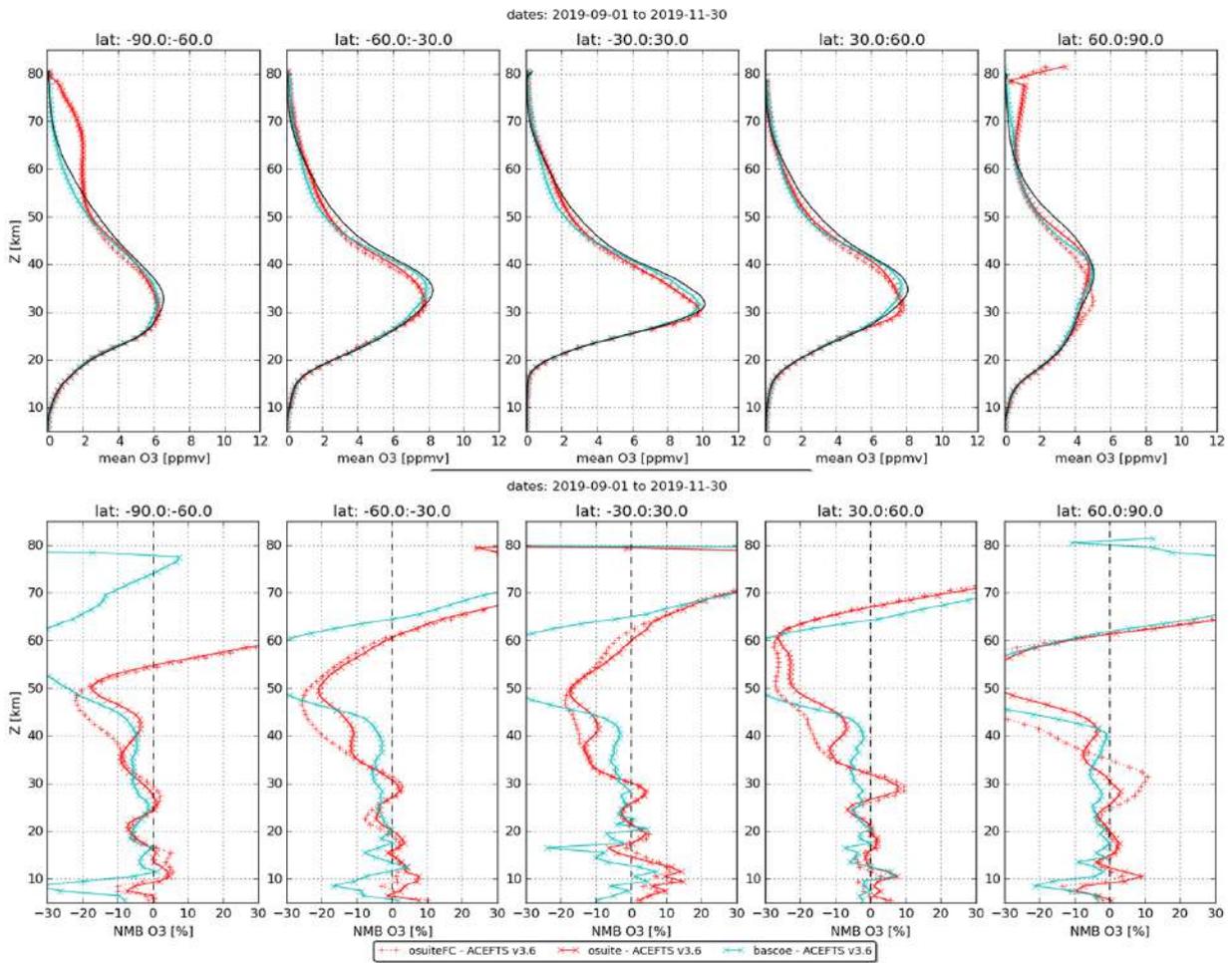
Figure 8.3.5: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with ACE-FTS observations for the period SON 2019.
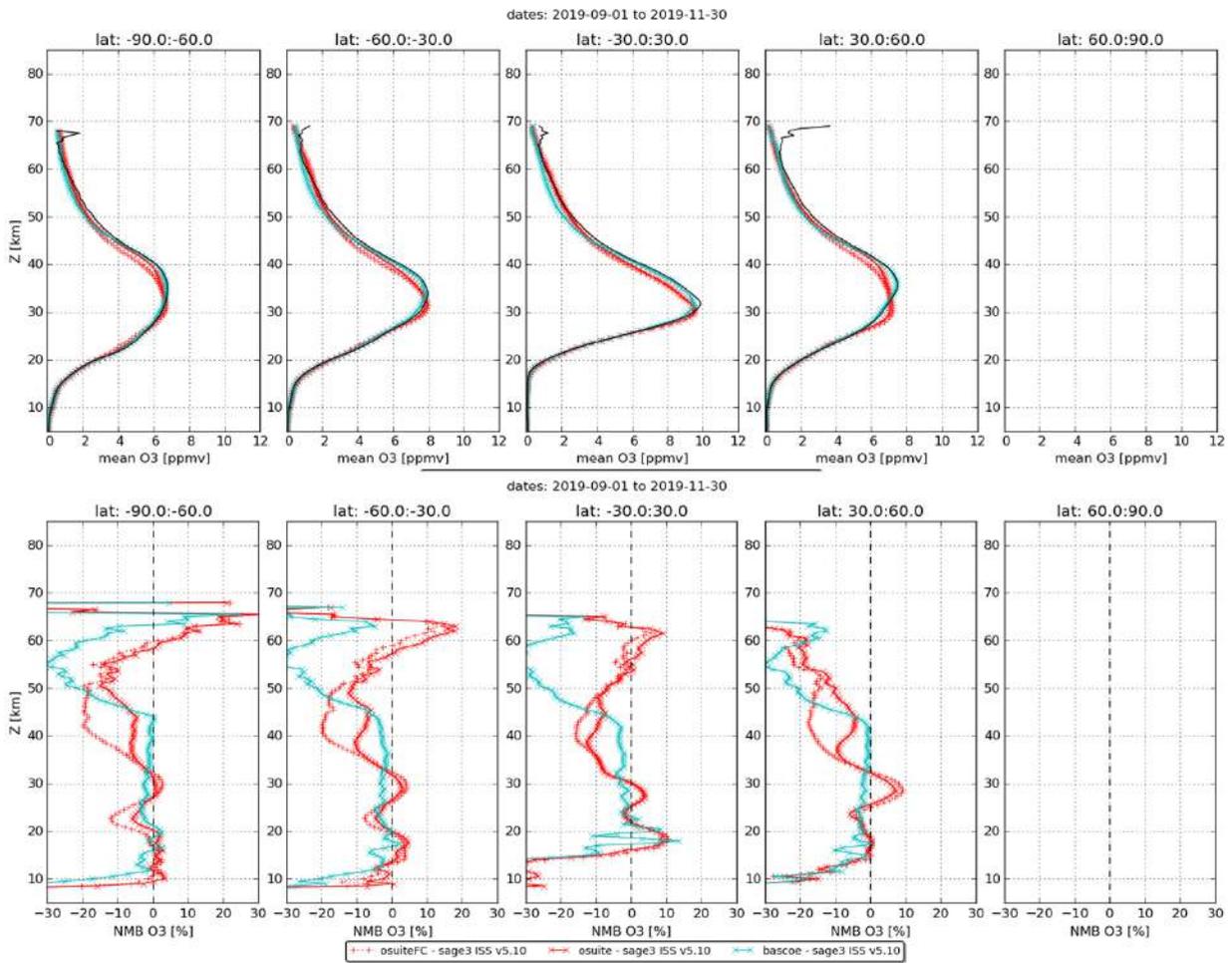
Figure 8.3.6: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with SAGE-III observations for the period SON 2019.
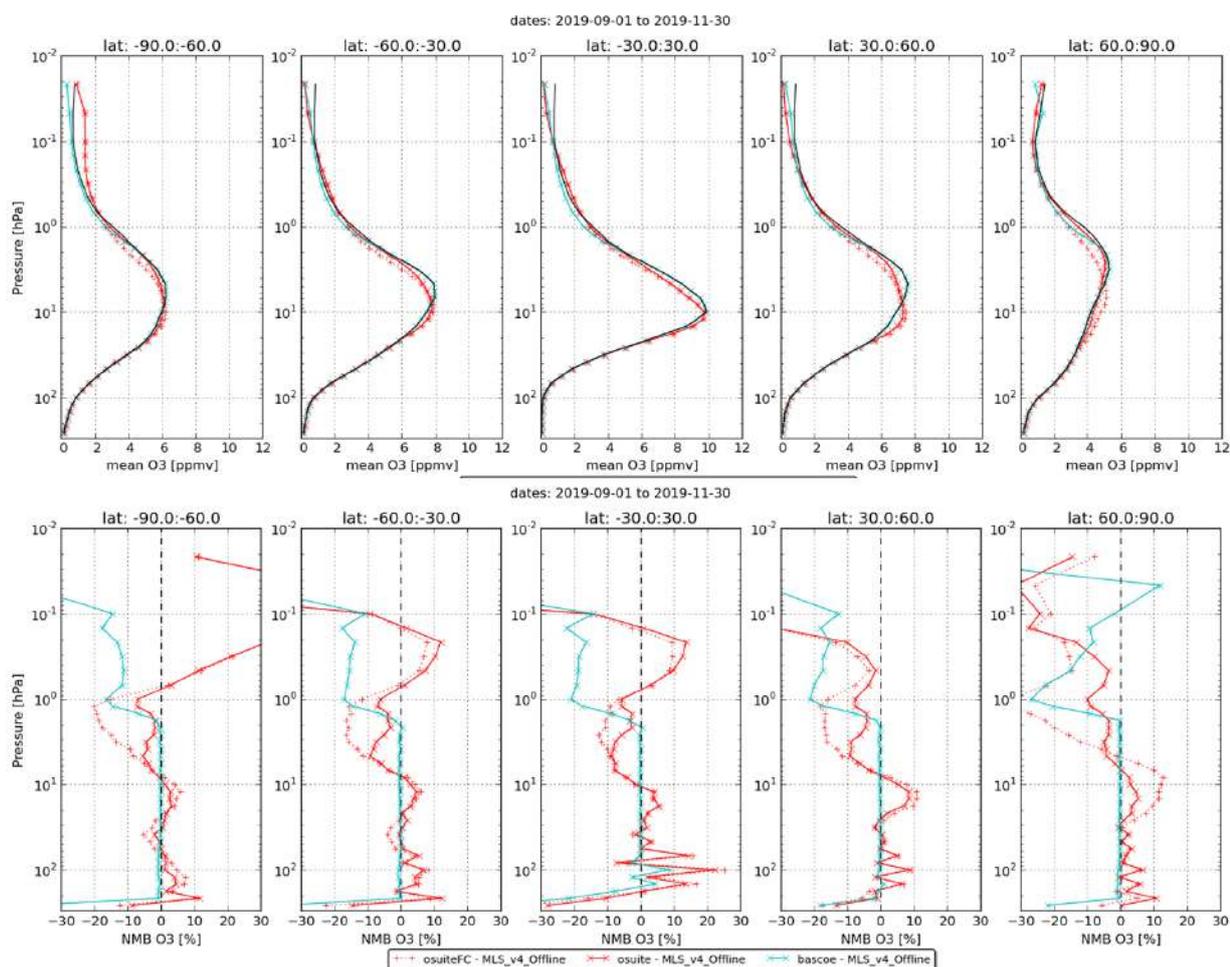
Figure 8.3.7: Mean value (top) and normalized mean bias (bottom) of the ozone profile between o-suite analyses (red, solid), o-suite forecasts 4th day (red, dotted) and BASCOE (cyan line) with MLS observations for the period SON 2019.

8.4 Stratospheric NO$_2$

The CAMS model uses a tropospheric chemistry scheme in combination with a parameterization for stratospheric ozone. Stratospheric ozone is also well constrained by satellite observations. Therefore, the only useful product in the stratosphere is ozone, and all other compounds, including NO$_2$, should not be used, as demonstrated by the validation results presented here.

In this section, nitrogen dioxide from SCIAMACHY/Envisat satellite retrievals (IUP-UB v0.7) and GOME-2/MetOp-A satellite retrievals (IUP-UB v1.0) are compared to modelled stratospheric NO$_2$ columns. Monthly mean stratospheric NO$_2$ columns from SCIAMACHY and GOME-2 have relatively small errors on the order of 20% in the tropics and in mid-latitudes in summer and even lower errors at mid-latitudes in winter. As the time resolution of the saved model files is rather coarse and NO$_x$ photochemistry in the stratosphere has a large impact on the NO$_2$ columns at low sun, some uncertainty is introduced by the time interpolation at high latitudes in winter.

As shown in Figure 8.4.1, amplitude and seasonality of satellite stratospheric NO_2 columns are poorly modelled with CB05-based chemistry runs including the more recent versions of the o-suite. The significant differences between observations and CB05 chemistry runs, i.e. a strong underestimation of satellite retrievals by models, can be explained by the missing stratospheric chemistry for these model versions. The only constraint on stratospheric NO_x is implicitly made by fixing the HNO_3/O_3 ratio at the 10 hPa level. This assumption, in combination with the changing model settings for stratospheric O_3 for control compared to MACC_CIFS_TM5, may explain some of the jumps we see in stratospheric NO_2. In any of these runs the stratospheric NO_2 is poorly constrained. It clearly indicates that stratospheric NO_2 in the latest versions of the o-suite is not a useful product and should be disregarded.

Comparison of the o-suite from July 2012 until August 2014 with the other model runs and satellite observations shows that the previous version of the o-suite stratospheric NO_2 columns had a systematic low bias relative to those from the Mozart-model based MACC system (MACC_fcnrt_MOZ) and satellite observations for all latitude bands. For example, o-suite values are a factor of 2 smaller than satellite values between 60°S to 90°S for October 2013. Best performance was achieved with the MOZART chemistry experiments without data assimilation (MACC_fcnrt_MOZ, running until September 2014), especially northwards of 30°S. Details on the NO_2 evaluation can be found at:
http://www.doas-bremen.de/macc/macc_veri_iup_home.html.
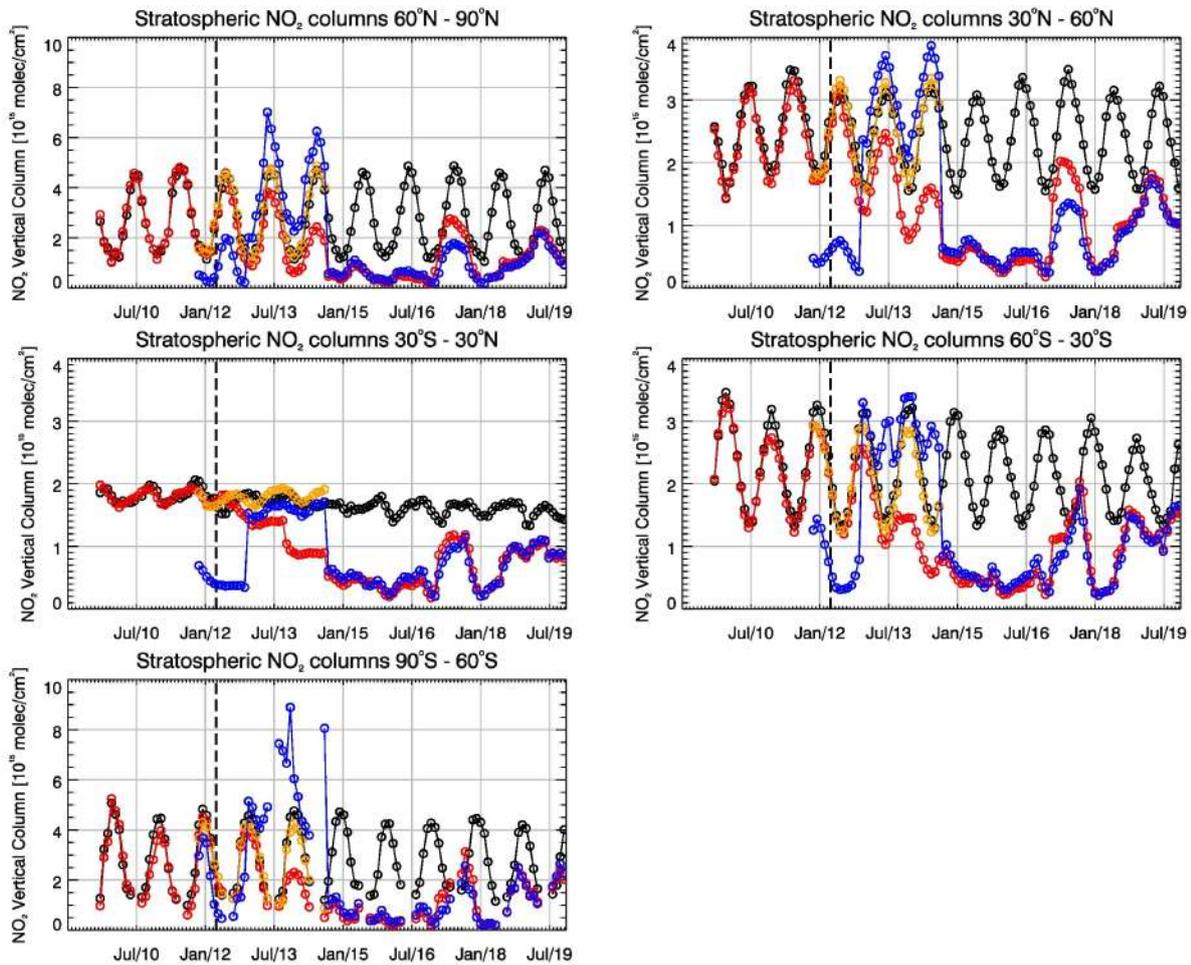
Figure 8.4.1: Time series of average stratospheric NO2 columns [10^{15} molec cm-2] from SCIAMACHY (up to March 2012) and GOME-2 (from April 2012, black) compared to model results (red: o-suite, blue: MACC forecast based on TM5 / MACC CIFS TM5 / control, orange: MACC forecast based on the Mozart model) for different latitude bands. See text for details. The blue line shows MACC_fcnrt_TM5 from November 2011 to November 2012, MACC_CIFS_TM5 results from December 2012 until August 2014 and control results from September 2014 onwards (the model run without data assimilation is termed control since Sep 2014). The vertical dashed black lines mark the change from SCIAMACHY to GOME-2 based comparisons in April 2012.

9. Validation results for greenhouse gases

This section describes the NRT validation of the pre-operational, high resolution forecast of CO_2 and CH_4 from 1st December 2018 to 1st December 2019 based on observations from 17 surface stations, located in Western Europe; 10 TCCON stations measuring XCO_2 and XCH_4 total columns, and 13 NDACC stations measuring partial and total CH_4 columns. We compare the observations to the high-resolution forecast experiments (*gqpe/gznv, Tco1279L137; 9x9 km*), coupled to the analysis experiment (*gqiq/gwx3, Tco399L137, 25x25 km*). The *gqpe* forecast experiment is using the IFS model cycle CY43R1, and has been officially implemented on 1st Nov. 2017. The *gznv/gwx3* experiments, based on IFS CY45R1, are used from 1st December 2018 on.

9.1 CH4 and CO2 validation against ICOS observations

The CO_2 and CH_4 simulations from the analysis and high-resolution forecast have been compared to the 17 ICOS stations. The near–real time data processing of the in-situ measurements is ensured by the Atmospheric Thematic Center (Hazan et al., 2016). Among the 17 stations we can distinguish three sites located on top of mountains (PUY, JFJ, CMN), two background sites (PAL, ZEP) and 12 tall towers. For the later we consider only in this report the highest sampling levels which are at least at 100m above the ground.

The figure 9.1.1 shows the time varying biases (observations minus model), averaged on a weekly basis, for all ICOS stations. The CO_2 biases display clear seasonal cycles (±10 ppm) at most sites with positive biases in Summer/Autumn, and negative biases in Winter/Spring. There is one clear exception at Ispra (IPR), a tall tower located in the Po valley, and poorly represented by the model due to the complex orography. It should also be noted that we observe a less negative bias at the mountain sites (JFJ, PUY, CMN), probably due to the fact that they are less exposed to surface regional fluxes. Two examples are detailed on Figure 9.1.2 for Jungfrauhoch (Switzerland) and Pallas (Finland). The first one is typical of the mountain sites, with an overestimation of the CO2 concentration, slowly increasing from 0 to 2% in from September 2018 to September 2019, and then decreasing down to 1% in December 2019. Pallas is representative of most ICOS stations. We observe the same range of increase of the bias from September 2018 to August 2019, from -1% to +2%. Then in mid-August there is an abrupt decrease of the bias (down to 0.5%). The same feature is observed at Nordic sites SMR, NOR, and one month later at most German and French sites (Figure 9.1.1). This change of bias corresponds to a shift in the CO2 minimum which is delayed in the model (see PAL, Figure 9.1.2). The increase in the positive bias is clearly seen in figure 9.1.4a showing the seasonal pattern of the model / observation differences averaged over 15 sites. The maximum European bias increases from 0 to 6-8 ppm from September 2018 to September 2019, and decreases in October. It should also be noted that the coefficient correlations are significantly lower and RMS higher from April to September, during the growing season of the vegetation when the biospheric fluxes are maximum.

For CH_4 the figure 9.1.1 shows two distinct patterns in the model-observations differences: a seasonal and a latitudinal one. At the Scandinavian sites the model overestimates (up to 50 ppb) the observations almost all along the year. The example of Norunda (Figure 9.1.3) clearly shows that
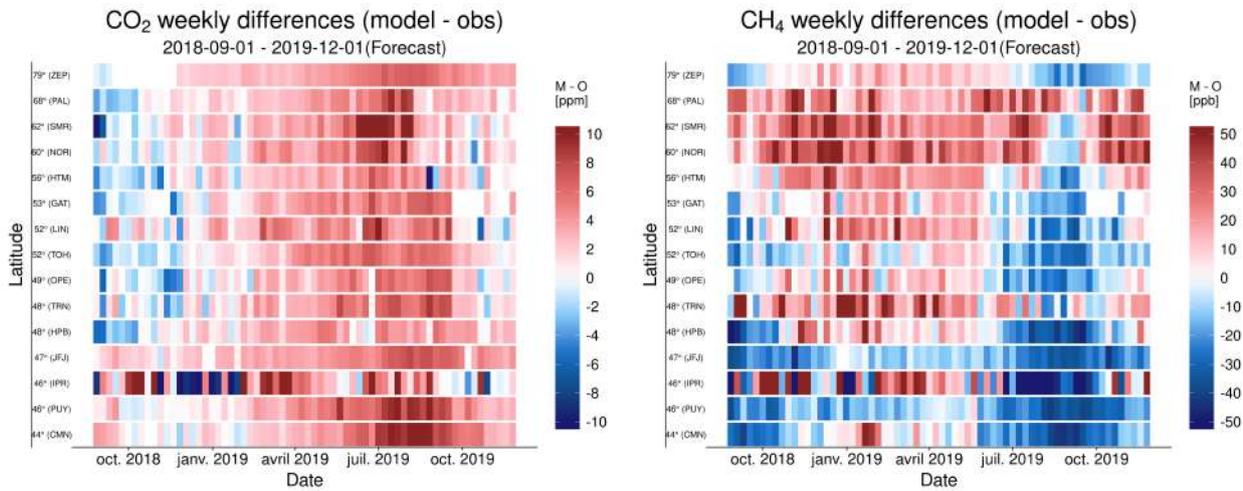
Figure 9.1.1: Mosaic plot of CO_2 (left, in ppm) and CH_4 (right, in ppb) biases of the CAMS high resolution forecast, compared to surface station observations. Each vertical colored line represents a weekly mean.
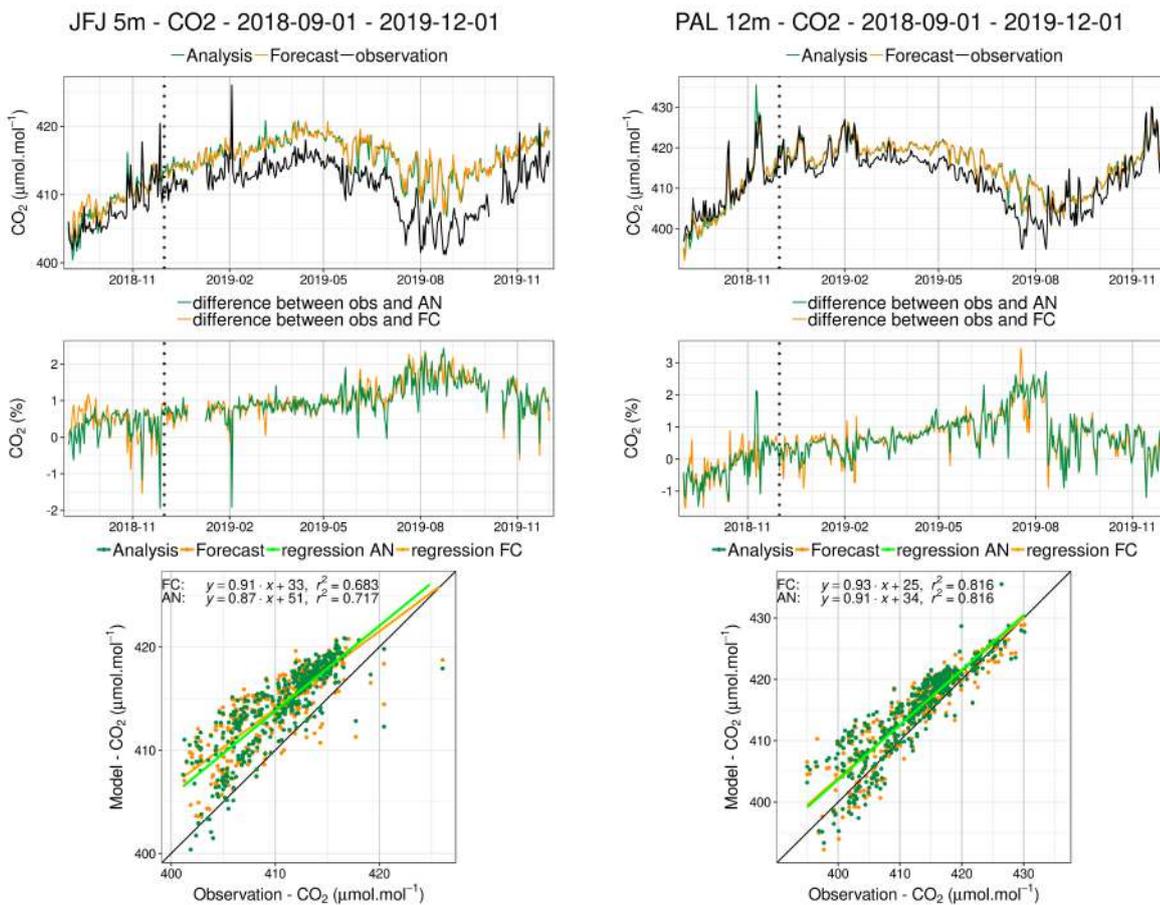

Figure 9.1.2: Comparison of CO_2 daily means observed (black) with the analysis run (green) and the high-resolution forecast (orange) at Jungfraujoch (left) and Pallas (right) for the period September 2018 to November 2019. Middle: differences of the observations minus the simulations. Below: Linear fit between observations and simulations. The dashed vertical line represents the change of experiments in December 2018.
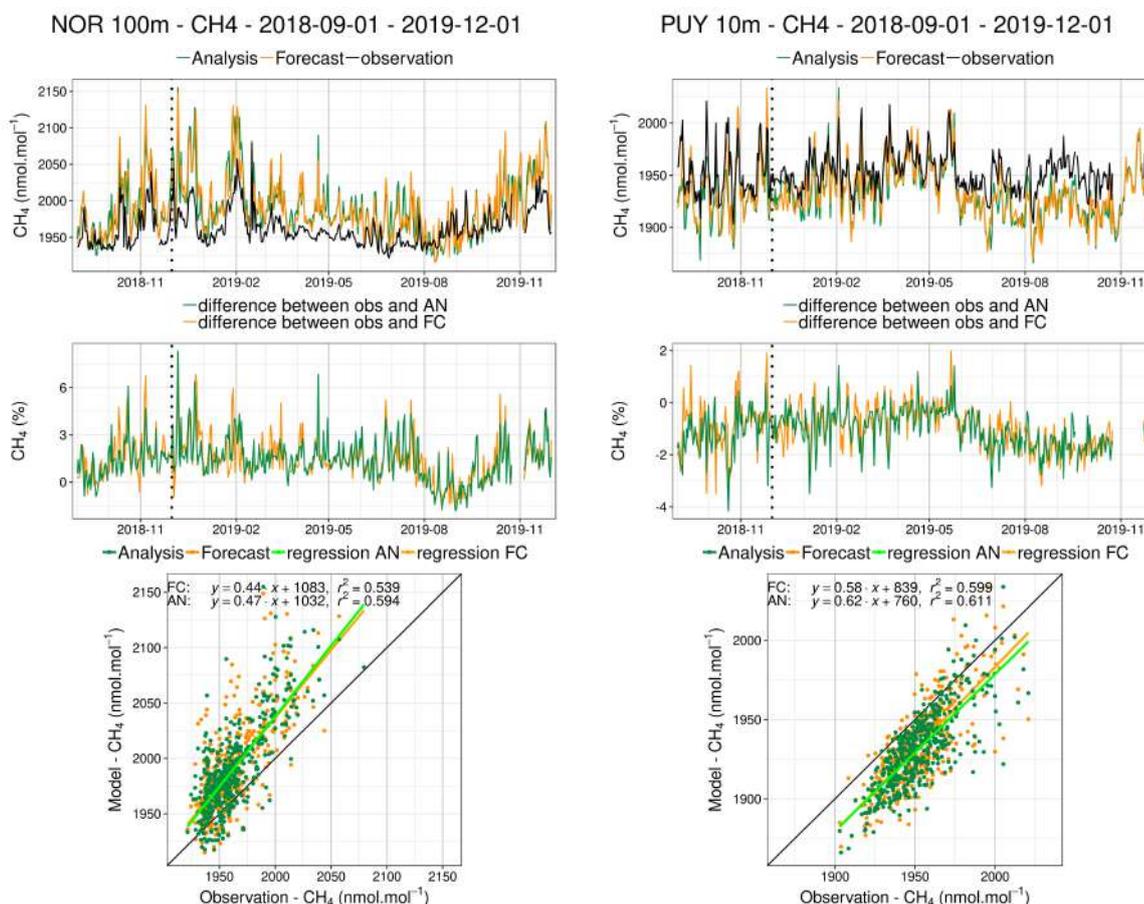
Figure 9.1.3: Comparison of CH4 daily means observed (black) with the analysis run (green) and the high-resolution forecast (orange) at Jungfraujoch (left) and Pallas (right) for the period September 2018 to November 2019. Middle: differences of the observations minus the simulations. Below: Linear fit between observations and simulations. The dashed vertical line represents the change of experiments in December 2018.

both the baseline and the CH4 spikes are overestimated by the model, which could indicate that the wetland emissions are overestimated. There is only a period in August/September 2019 when the observations are greater than the simulated values (Figure 9.1.3). The more we go to lower latitudes, the more we observe negative biases. In Germany and Northern France the bias is negative in Summer/Autumn. Negative biases are observed all year long for the mountain sits of Puy de dome. When considering the average signal from 15 European sites (Figure 9.1.4b), we observe a mean positive bias up 20 ppb until May 2019, and a small negative bias (about -5 ppb) from June to December 2019. However, considering the high latitudinal gradient (Figures 9.1.1), the mean CH4 European bias must be taken with extreme caution.

Analysis - Monthly metrics in the northern hemisphere
CMN-GAT-HPB-HTM-IPR-JFJ-LIN-NOR-OPE-PAL-PUY-SMR-TOH-TRN-ZEP
CO_2 ($\mu mol.mol^{-1}$)
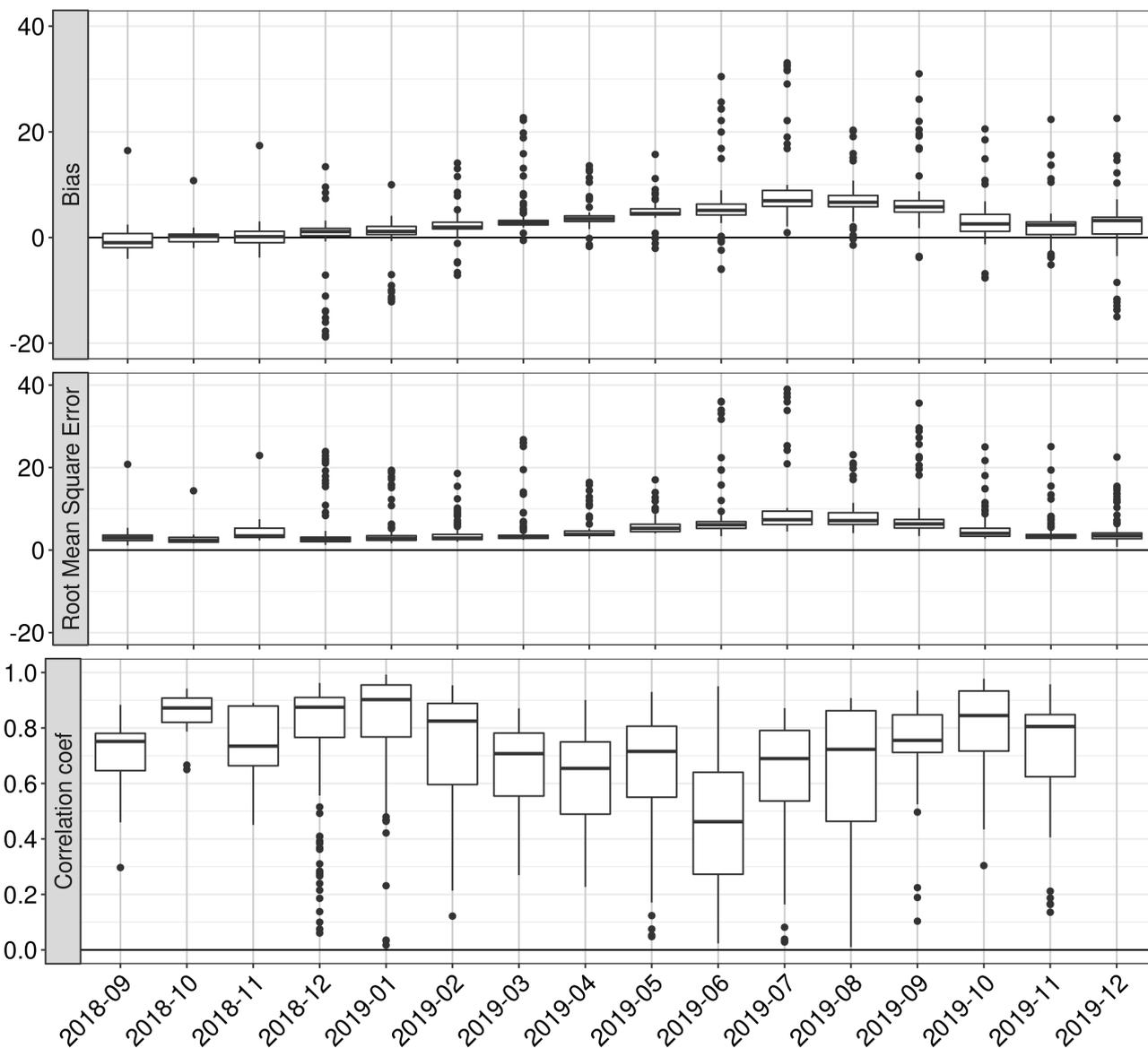
Figure 9.1.4a: Monthly statistics (bias, RMSE, correlation coefficients) of the analysis experiment compared to CO_2 surface measurements at ICOS sites. The results obtained for all European sites (see the list of sites in the title) are averaged. September 2019 is not representative, since using only one day.

Analysis - Monthly metrics in the northern hemisphere
CMN-GAT-HPB-HTM-IPR-JFJ-LIN-NOR-OPE-PAL-PUY-SMR-TOH-TRN-ZEP
CH_4 $(nmol.mol^{-1})$
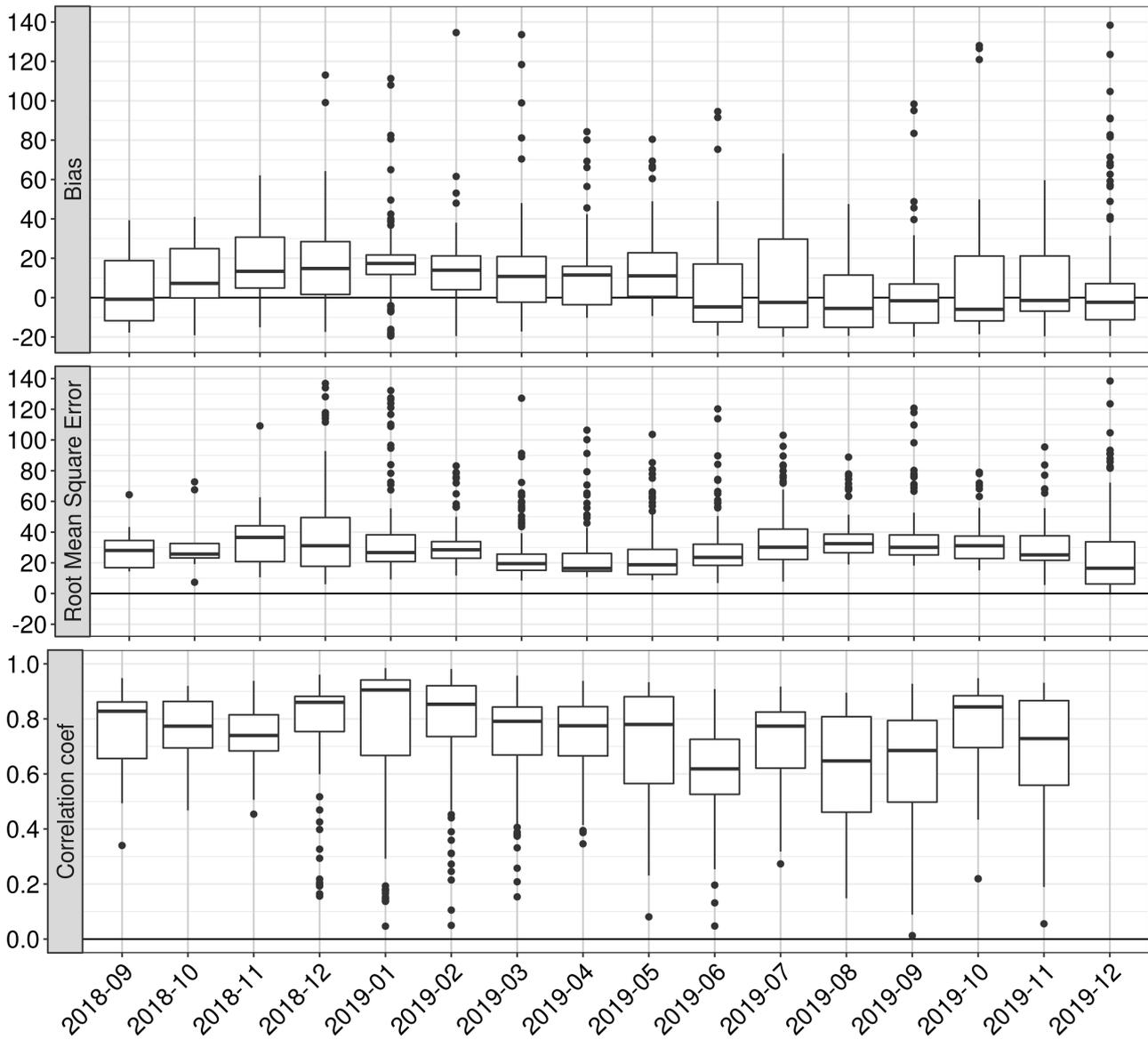
Figure 9.1.4b: Same as Figure 9.1.4a for CH_4.

9.2 CH₄ and CO₂ validation against TCCON observations

For the validation column averaged mole fractions of CO_2 and CH_4 (denoted as XCO_2 and XCH_4) from the Total Carbon Column Observing Network (TCCON) are used. Column averaged mole fractions provide different information than the in-situ measurements and are therefore complementary to the in-situ data.

The validation routines used for TCCON data are the same as used for the NDACC network and are documented in Langerock et al. (2015). The routines have been adapted to use the TCCON data format. In this section, we compare column averaged mole fractions of CO_2 and CH_4 of the CAMS models with TCCON retrievals. Data from the following TCCON sites has been used:
Izana (Blumenstock et al., 2017), Reunion (De Mazière et al., 2017), Bialystok (Deutscher et al., 2017), Manaus (Dubey et al., 2017), Four Corners (Dubey et al., 2017), Ascension (Feist et al., 2017), Anmeyondo (Goo et al., 2017), Darwin (Griffith et al., 2017), Wollongong (Griffith et al., 2017), Karlsruhe (Hase et al., 2017), Edwards (Iraci et al., 2017), Indianapolis (Iraci et al., 2017), Saga (Kawakami et al., 2017), Sodankyla (Kivi et al., 2017), Hefei (Liu et al., 2018), Tsukuba (Morino et al., 2017), Burgos (Morino et al., 2018), Rikubetsu (Morino et al., 2017), Bremen (Notholt et al., 2017), Spitsbergen (Notholt et al., 2017), Lauder (Sherlock et al., 2017, Pollard et al., 2019), Eureka (Strong et al., 2018), Garmisch (Sussmann et al., 2017), Zugspitze (Sussmann et al., 2018), Paris (Te et al., 2017), Orleans (Warneke et al., 2017), Park Falls (Wennberg et al., 2017), Caltech (Wennberg et al., 2017), Lamont (Wennberg et al., 2017), Jet Propulsion Laboratory (Wennberg et al., 2017), East Trout Lake (Wunch et al., 2017)

For the validation of the models in September, October and November data from 10 TCCON sites were available. For the sites Orleans and Reunion data was available for the full reporting period. Data from a new site in Nicosia (Cyprus) is also available but could not be included in time due to technical problems. It will be included in the next report.

Methane (CH₄)

Figure 9.2.1 shows the data for the period March 2018 to November 2019 for the high-resolution greenhouse gas model runs (top panel) and the greenhouse gas analysis in the lower panel. The upgrade of the NRT GHG analysis and high-resolution forecast to CY45R1 on 1 December 2018 are indicated.

The figure shows that both CAMS GHG configurations underestimate the CH_4 for all stations, with negative biases between 0 and 25 ppb, except for the sites around Los Angeles (Pasadena and JPL). The largest biases occur in July-November, with smaller biases in winter-spring. The mean biases are shown in Fig. 9.2.4 for the reporting period SON-2019.

The Taylor diagrams are provided in Fig. 9.2.2. Correlations range between 0.45 and 0.92 and the high-resolution forecast shows a higher standard deviation than the analysis. The time series at Orleans is shown in Fig. 9.2.3, again showing the relatively large bias in the Autumn of 2019.
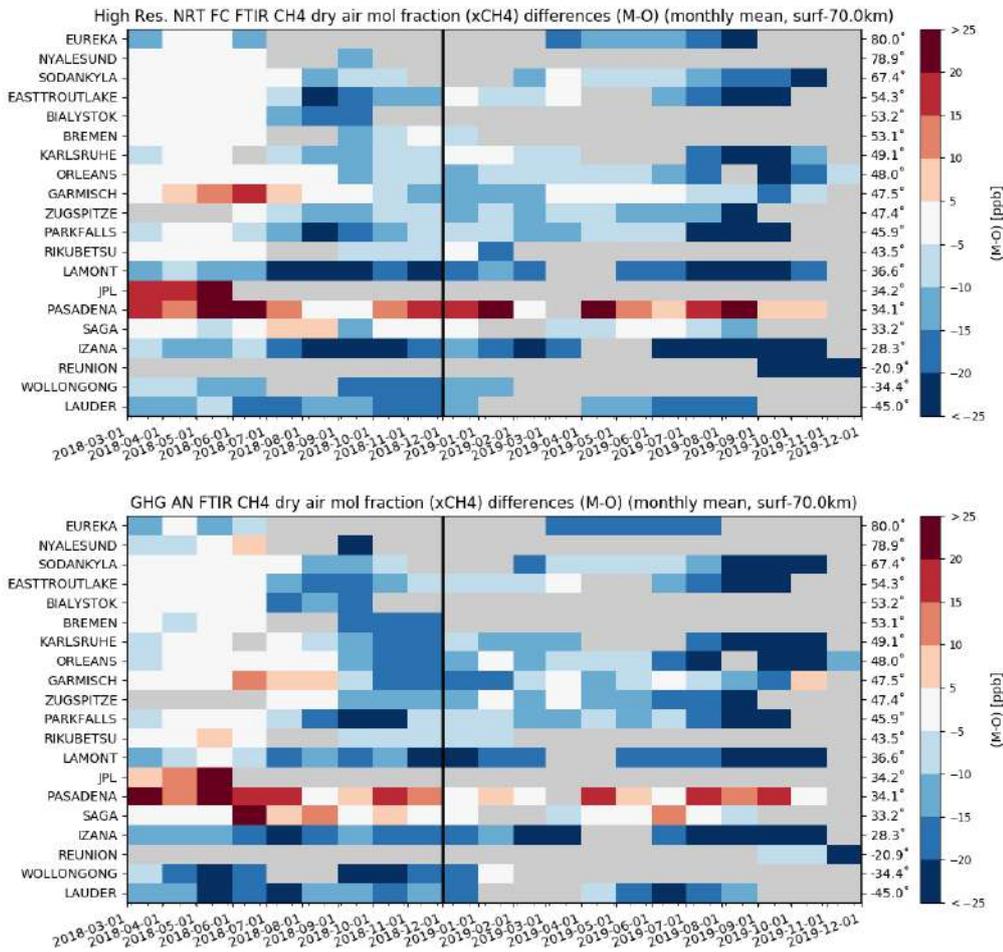
Figure 9.2.1: Monthly differences since March 2018 (upper plot: high-resolution NRT, lower plot: greenhouse gas analysis). The stations are sorted by latitude (northern to southern hemisphere).
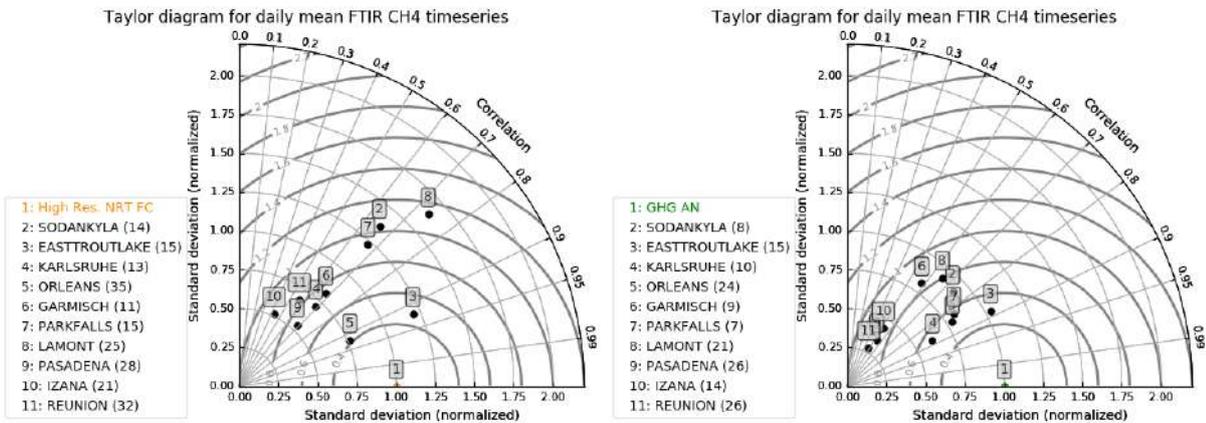
Figure 9.2.2: Taylor diagrams for the reporting period September-November 2019 (left: high-resolution near-real time; right: greenhouse-gas analysis).
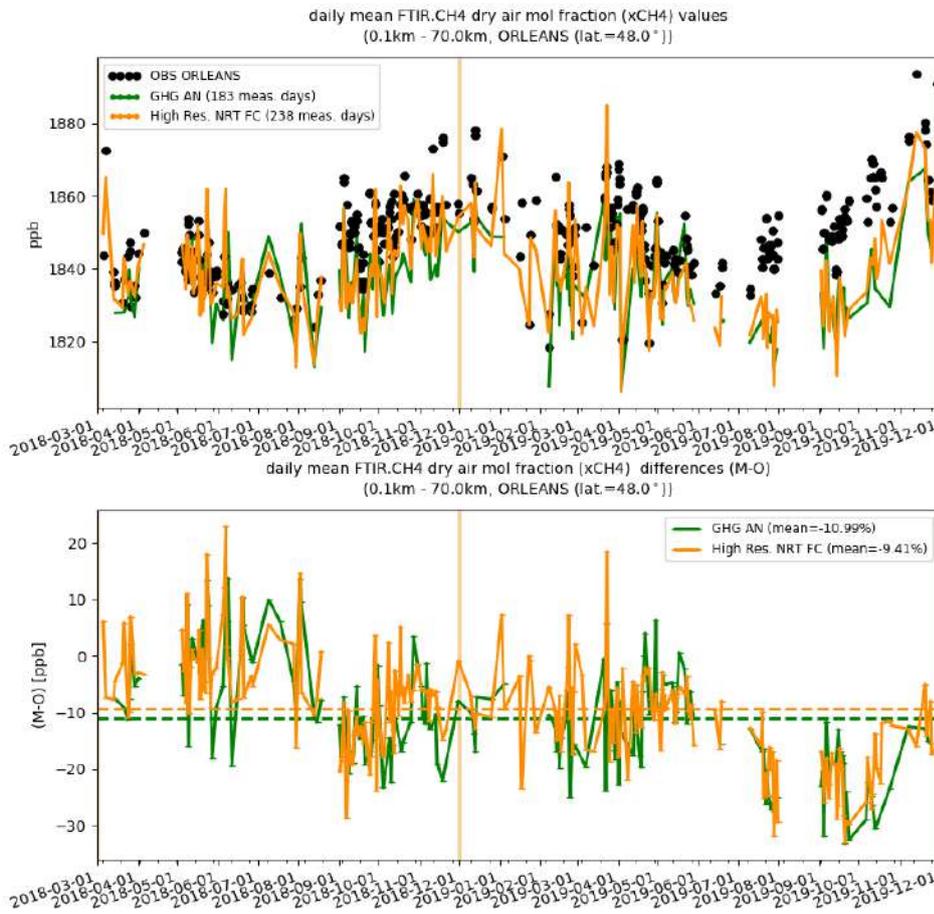
Figure 9.2.3: Comparison of the XCH$_4$ model data with TCCON data at Orleans.
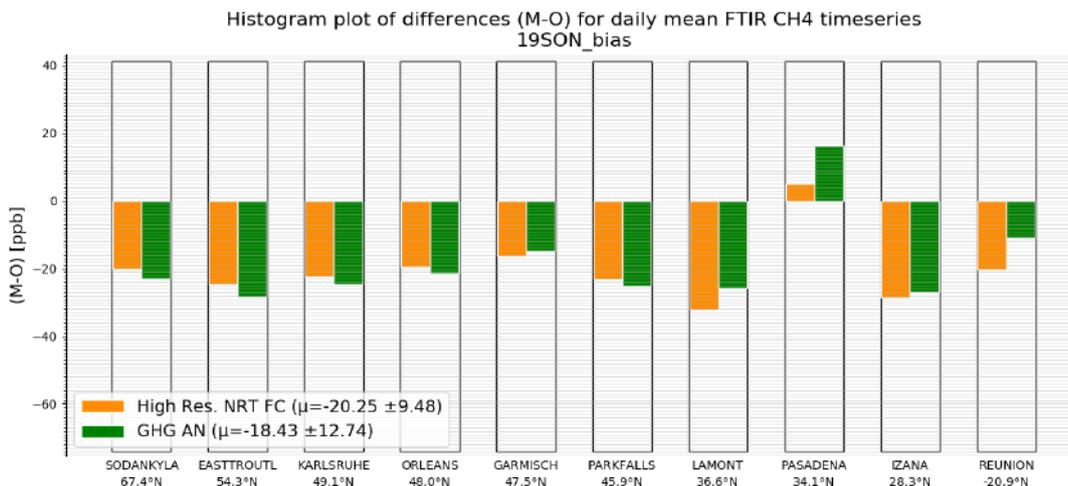
Figure 9.2.4: Differences during the reporting period

Carbon dioxide (CO$_2$)

Figure 9.2.5 shows the data for the last 4 years. The data show that the model data strongly overestimates the CO$_2$, with biases between 0 and 10 ppm. The largest biases occur in the Autumn.
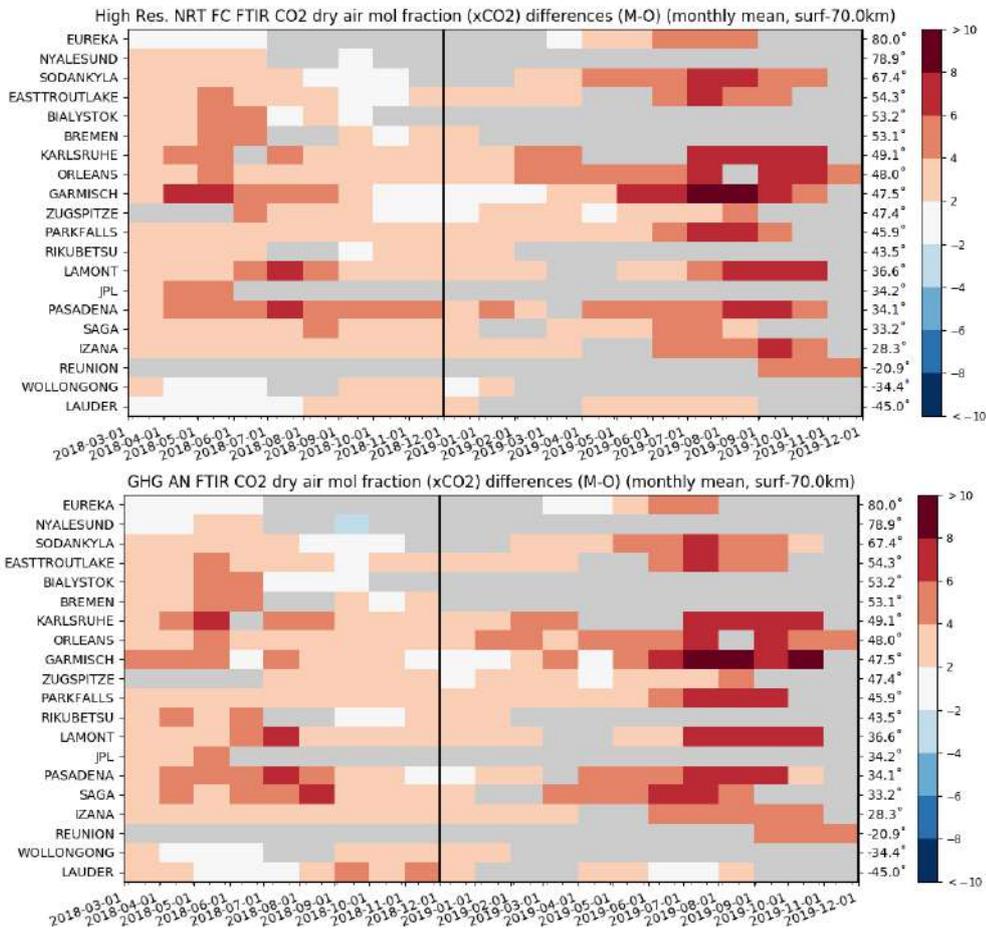
Figure 9.2.5: Monthly differences for the period since March 2018 (upper plot: high res NRT, lower plot: GHG AN). The stations are sorted by latitude (northern to southern hemisphere).
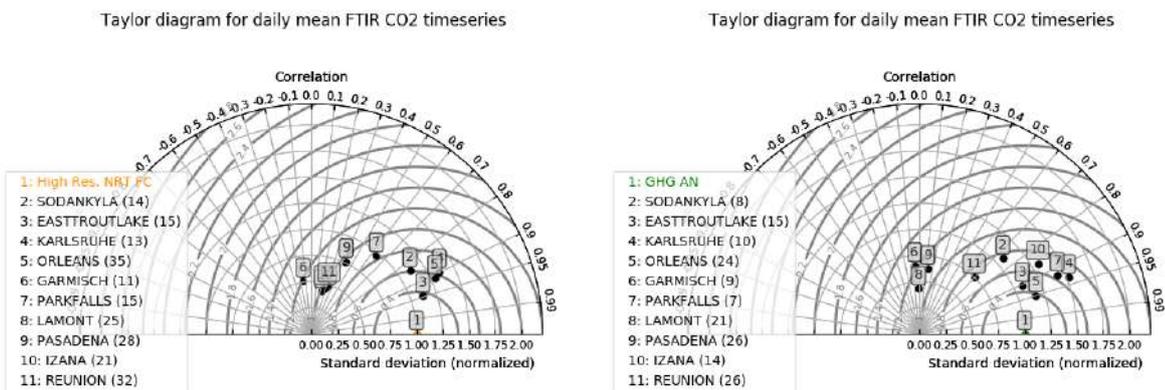
Figure 9.2.6: Taylor diagrams for the reporting period Sep-Nov 2019 (left: high res NRT, right: GHG AN).

The Taylor plots (Fig. 9.2.6) show that the forecast and analysis run behave similarly. The time series at Orleans, Fig. 9.2.7, shows the negative bias (see also Fig. 9.2.8), but also show that the seasonal cycle is tracked well by the CAMS configurations (high correlations in the Taylor plots).
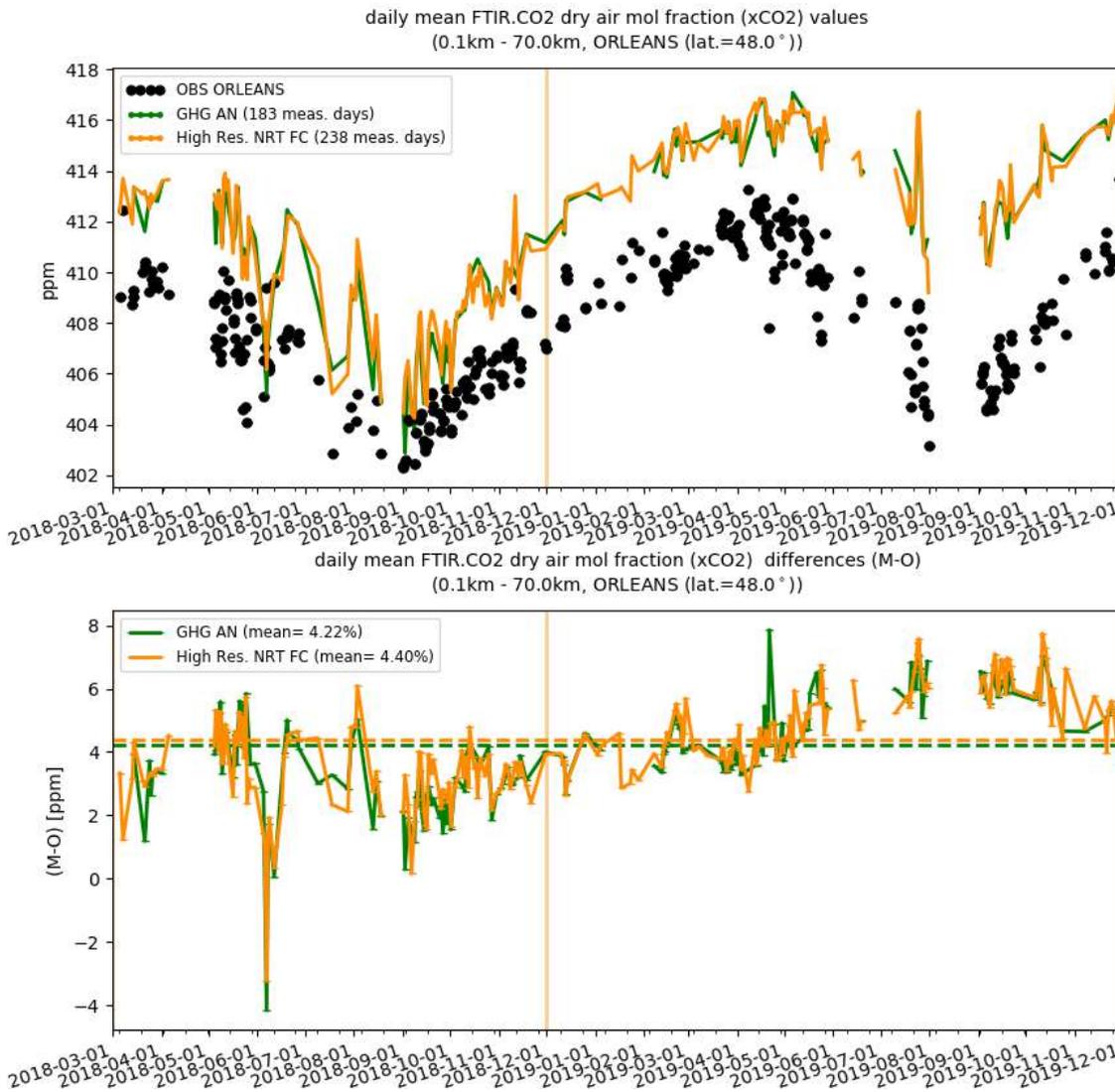
Figure 9.2.7: Comparison of the CO_2 model data with TCCON CO_2 at Orleans.
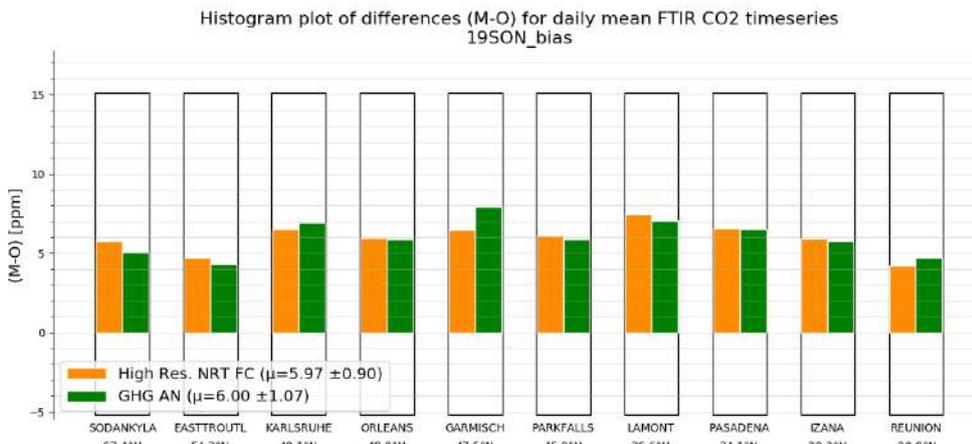
Figure 9.2.8: Differences during the reporting period
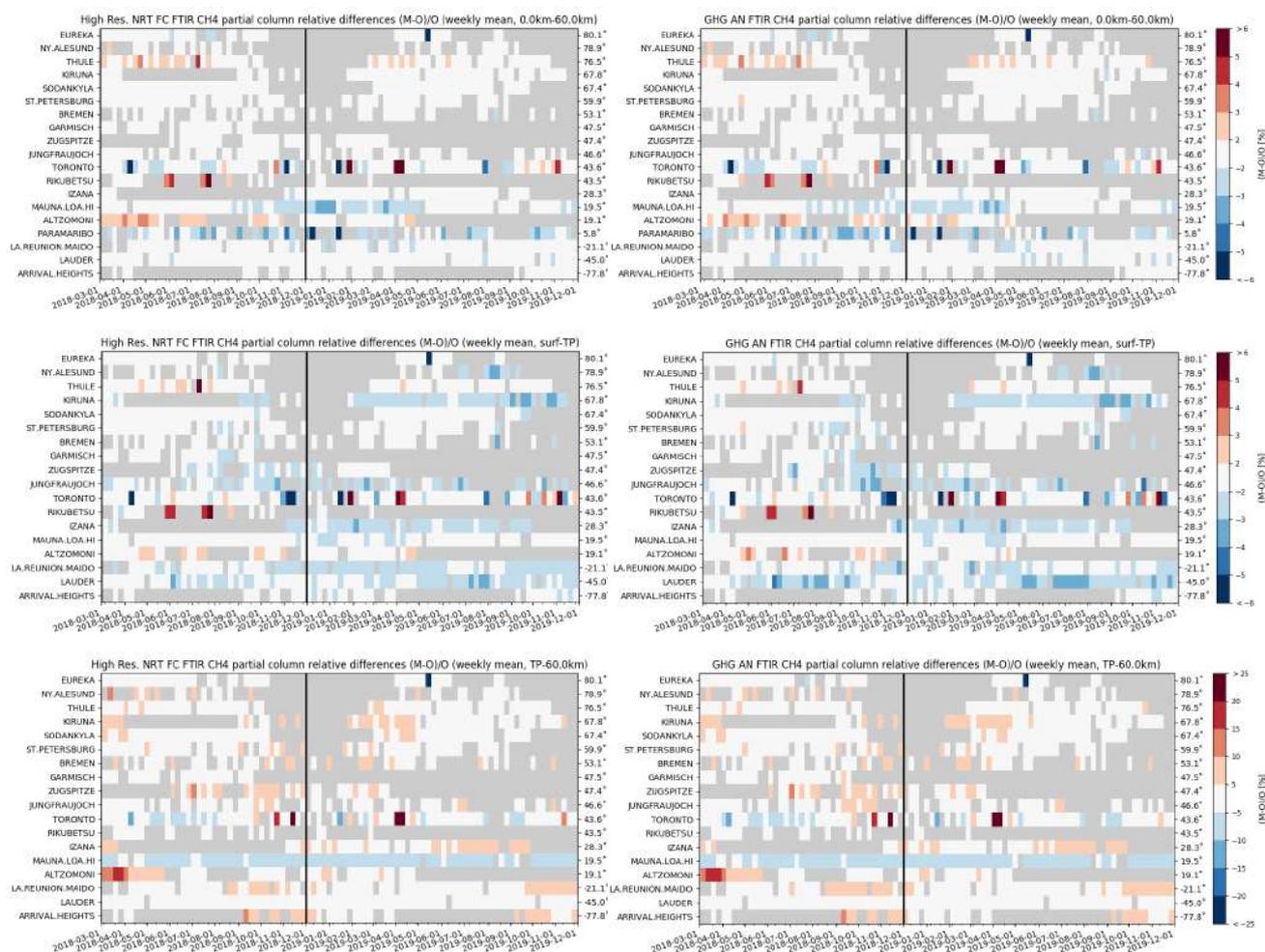
Figure 9.3.1: Weekly mean relative bias for total (top row), tropospheric (middle row) and stratospheric CH_4 columns (bottom row) for the period March 2018 – November 2019 for high resolution forecast (left) and the analysis (right). The overall uncertainty for the CH_4 total column measurements is approximately 4%. The overall uncertainty for the CH_4 total/tropospheric column measurements is approximately 2%, while the stratospheric uncertainty is 7.5% (color scale for the mosaic plots follows uncertainty scale)

9.3 Validation against FTIR observations from the NDACC network

In this section, we compare the CH_4 profiles of the CAMS GHG products with FTIR measurements at different FTIR stations within the NDACC network. These ground-based, remote-sensing instruments are sensitive to the CH_4 abundance in the troposphere and lower stratosphere, i.e. between the surface and up to 25 km altitude. Tropospheric and stratospheric CH_4 columns are calculated from the FTIR profile data and used to validate corresponding columns obtained from the model data. A description of the instruments and applied methodologies can be found at http://nors.aeronomie.be. The typical uncertainty on the FTIR tropospheric column is 2%, while the uncertainty on the stratospheric column is 7.5%, adding together to a 3% uncertainty on the total column. The systematic uncertainty is large for the NDACC methane product mostly due to higher spectroscopic uncertainties.
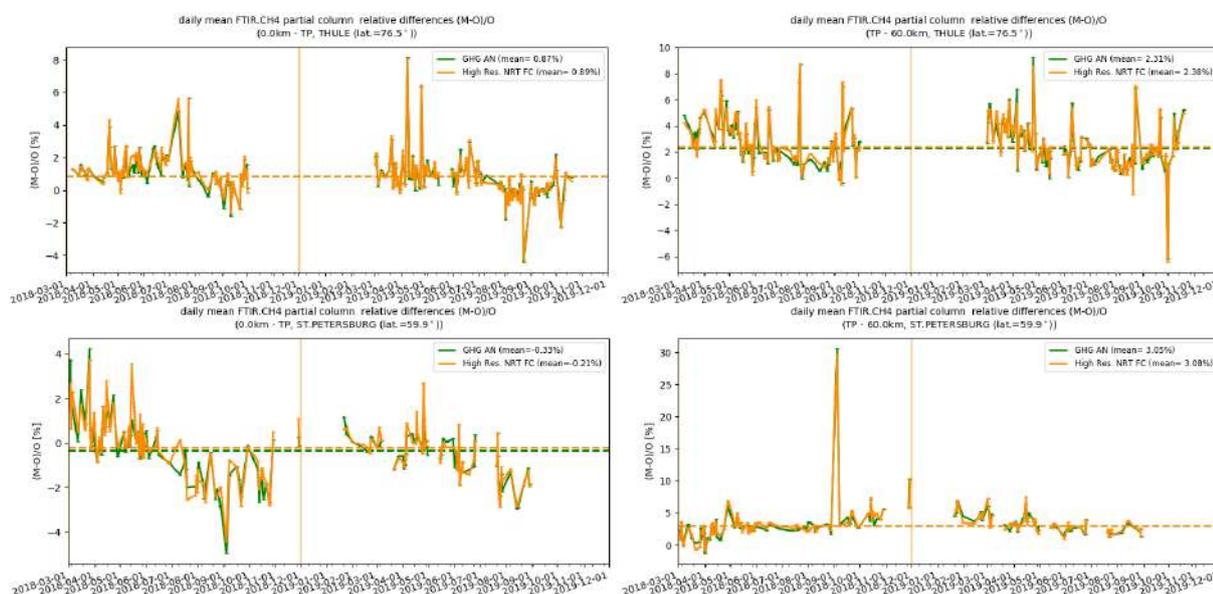
Figure 9.3.2: Daily mean of relative differences for tropospheric CH$_4$ columns (left) and stratospheric CH$_4$ columns (right) at Thule (top) and St Petersburg (bottom). At Thule the stratospheric column shows a reduced bias during the summer months, while at St Petersburg the tropospheric column performs worse during June-October.

Figure 9.3.1 (middle row) shows that the tropospheric columns of CH$_4$ agree well and only small differences appear between the analysis and the high-resolution model. In comparison with the measurement uncertainty, a slight underestimation is observed in the tropospheric columns which is in agreement with the TCCON results. The Paramaribo measurements have reduced sensitivity and the tropospheric/stratospheric split is not valid in this case.

The stratospheric columns (Figure 9.3.1, bottom row) show a slight overestimation compared to the measurement uncertainty.

At some sites a seasonal change is observed in either the tropospheric or stratospheric concentrations. Due to the short time period, it is unclear if this is a recurring seasonal dependent model performance. In Figure 9.3.2 the tropospheric and stratospheric relative difference time series are plotted at Thule and St. Petersburg.

Figure 9.3.3 shows Taylor diagrams for the SON time period and for a selected number of sites (many high latitude stations are not measuring during SON): some stations have limited observations and should be treated with care. Assimilation has a small effect on the correlation coefficients for most sites: the average correlation for 11 stations is 0.76 for the analysis and 0.84 for the high-resolution forecast.
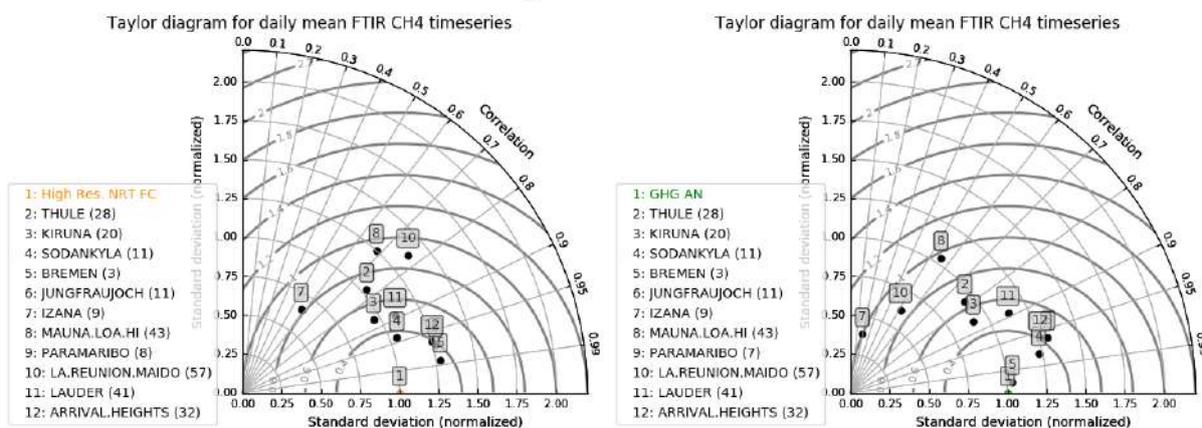
Figure 9.3.3: Taylor diagrams relating the standard deviations for the model /GB time series of total CH$_4$ column data and their correlation for the period 2019 SON (the stations with a limited number of measurements should be ignored). All time-series are normalized such that the std of the model column time series is 1. For the tropical sites Mauna Loa (8) and Reunion (10) the assimilation decreases the ratio of the standard deviations of both time series significantly: the analysis methane columns are more variable compared to the high-resolution forecast.

10.   Event studies

10.1 High surface ozone episodes in Europe during 26-30 June and 24-27 July 2019.

The CAMS NRT performance was validated for the two major heat waves over western and central Europe, during 26-30 June 2019 and 24-27 July 2019 against Airbase observations. Fig 10.1.1 shows the spatial distribution of the air temperature at 2m above the surface anomalies from the 1981-2010 JJA climatology, during 26-30 June 2019 (left) and 24-27 July 2019 (right), where it is evident that in a large part of western and central Europe air temperatures exceeds the JJA 1981-2100 climatological mean up to 8 ºC.

In both cases the heat waves lead to high ozone episodes. This is clearly seen in Figures 10.1.2 and 10.1.3 (left maps) where it is evident that observed O3 daily mean values and observed O3 daily maximum values during these periods are higher than the JJA 2019 mean up to 75% with the maximum exceedance observed over France, Switzerland and Belgium (and also over eastern UK for the 24-27 July 2019 episode). In Figures 10.1.2 and 10.1.3 centre and right graphs are plotted the same O3mean and O3max anomalies calculated for CAMS o-suite and CAMS control respectively. In these graphs we can see that the CAMS NRT simulations can reproduce the observed anomalies pattern but underestimates its magnitude.

This is also evident in Figure 10.1.4 where the observed and simulated surface ozone 3-hourly France stations mean time-series are plotted together with the CAMS air temperature values. In this graph is clear that although the CAMS NRT systems are able to reproduce the summer day to day variability they underestimate the peak values.
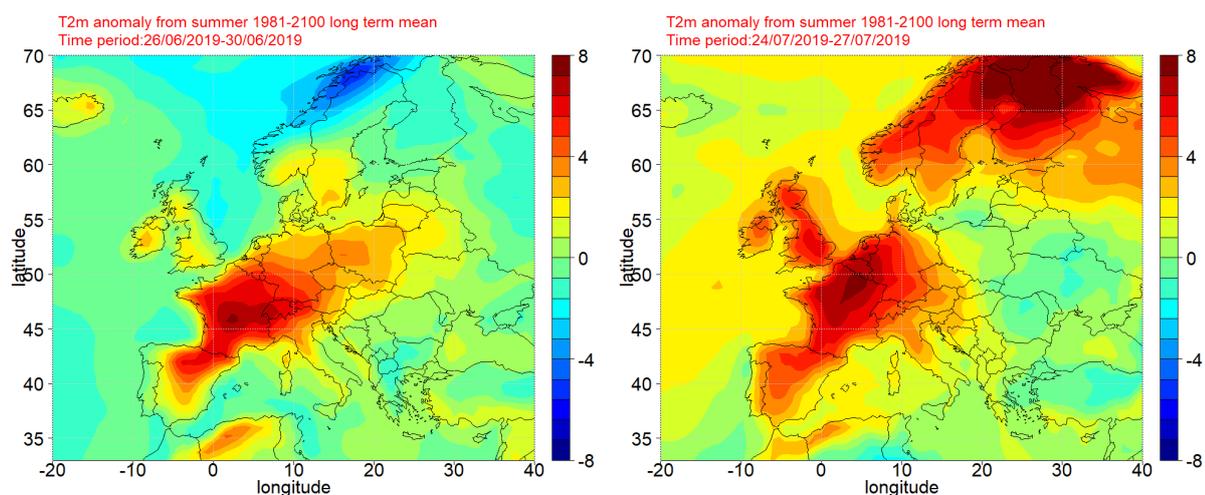
Figure 10.1.1. Spatial distribution of the air temperature anomaly at 2m above the surface (in ºC) during 26-30 June 2019 (left) and 24-27 July 2019 (right), from the 1981-2010 JJA climatology. Source of the data: ECMWF ERA-Interim reanalysis.
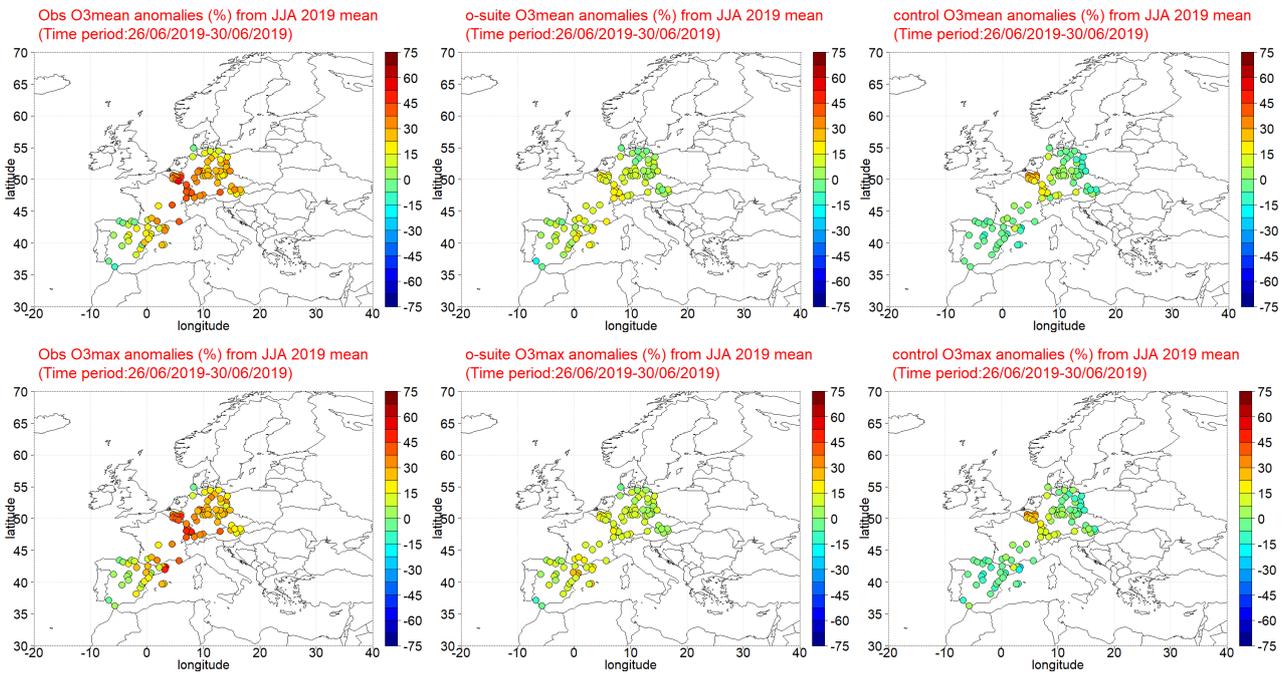
Figure 10.1.2. Spatial distribution of the surface ozone anomalies (in %) during 26-30 June 2019 from JJA 2019 mean for Airbase observations (left), o-suite (centre) and control run(right) simulated values. Upper graphs correspond to daily mean values and lower graphs to daily maximum values.
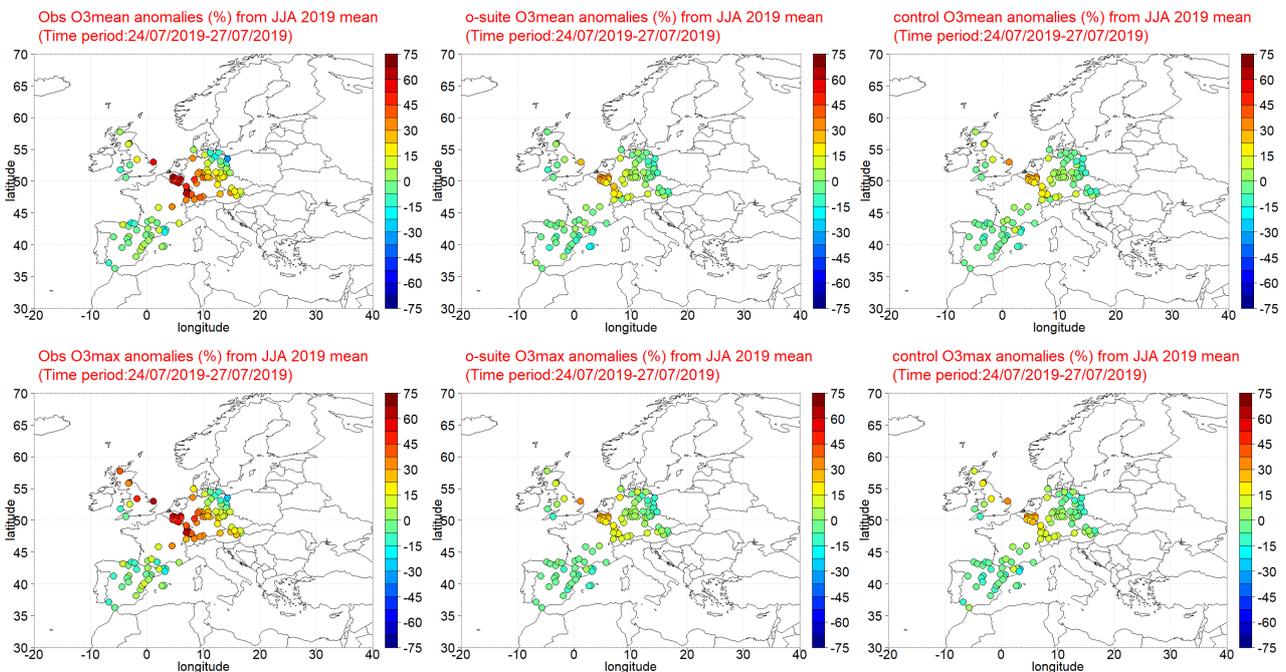
Figure 10.1.3. Spatial distribution of the surface ozone anomalies (in %) 24-27 July 2019 from JJA 2019 mean for Airbase observations (left), o-suite (centre) and control run(right) simulated values. Upper graphs correspond to daily mean values and lower graphs to daily maximum values.
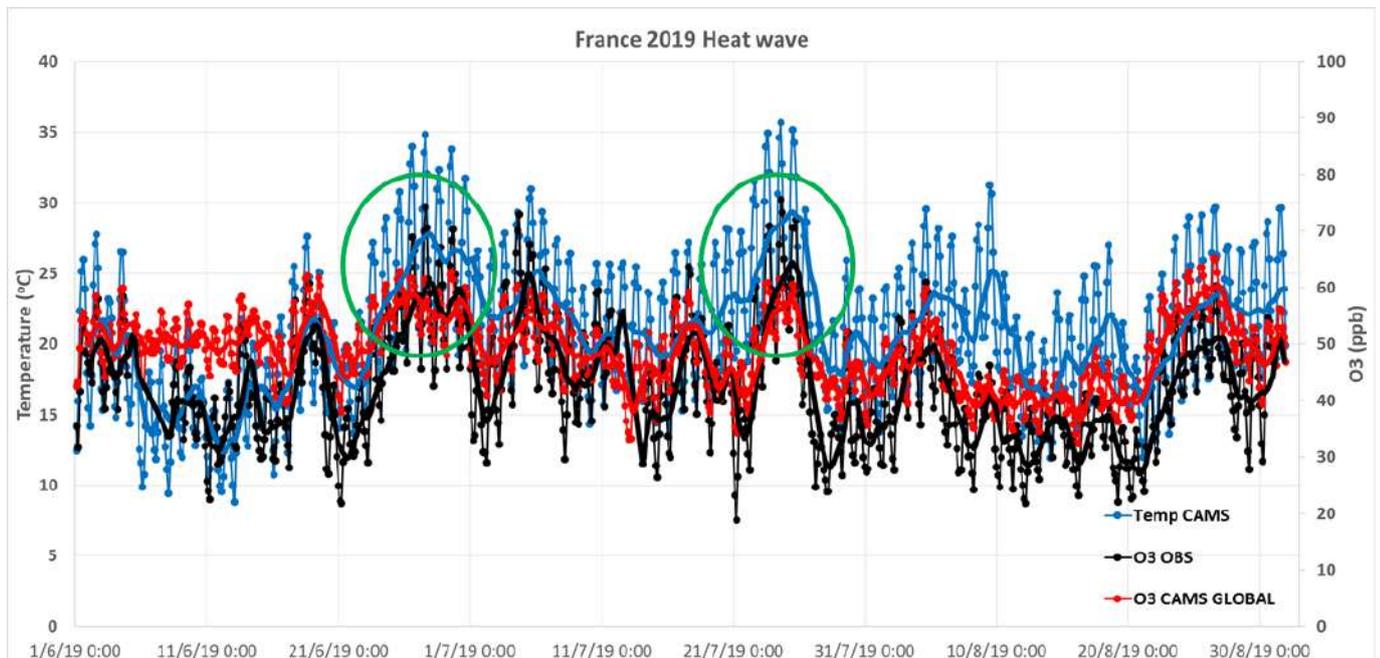
Figure 10.1.4: Observed (black) and CAMS o-suite (red) simulated surface ozone together with the CAMS air temperature (blue) 3hourly France stations mean time-series for the JJA 2019 time period. The green cycles denote the 26-30 June 2019 and 24-27 July 2019 high ozone episodes.

10.2 Iberian Peninsula dust event in early-September 2019

In mid-September 2019, MODIS satellite detected two dust outbreaks that reached the Iberian Peninsula (see Figure 10.2.1). One outbreak occurred on 13-17 September June, originating from Algeria and transported towards the Western Mediterranean, hitting the Balearic Islands and Southern France, and moving to the west crossing the Iberian Peninsula. The event was nicely tracked by MODIS (see Figure 10.2.1 and Figure 10.2.2) and in-situ observations.

CAMS AOD o-suite did timely reproduce the spatial distribution of the two dust plumes as shown the comparison with MODIS/Aqua AOD comparison (see Figure 10.2.2) despite the model tending to overestimate the observed maximum values. The Madrid and Zaragoza AERONET sites in Central Spain (Figure 10.2.3) show values of AOD up to 0.5 and the o-suite predicts AOD values around 0.6 on 15th September. During this event, the o-suite is predicting PM10 values over 300µg/m3 in large extensions of the Iberian Peninsula (Figure 10.2.2), i.e. overestimating the PM10 observations in Spain and Portugal (Figure 7.4.5).
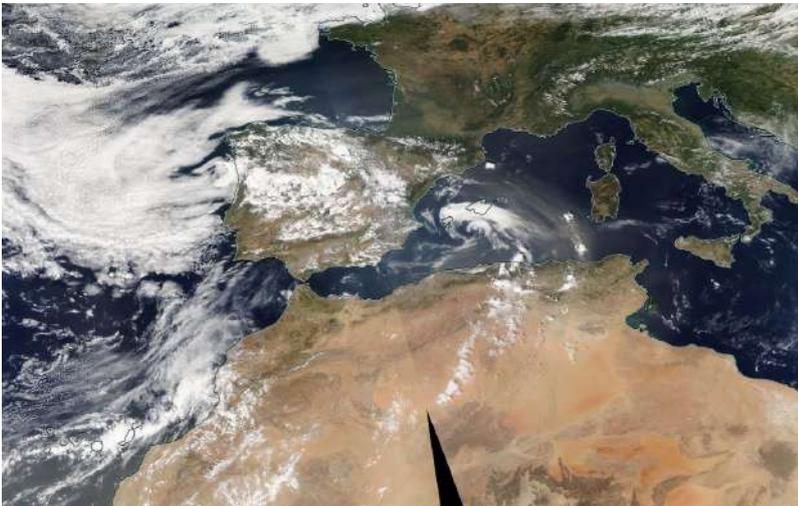
Figure 10.2.1: Daily AOD at 550nm composite of NASA MODIS Terra/Aqua on 16th September over Iberian Peninsula. These images are a zoom of the images included in the comparison with CAMS o-suite in Figure 10.2.2.
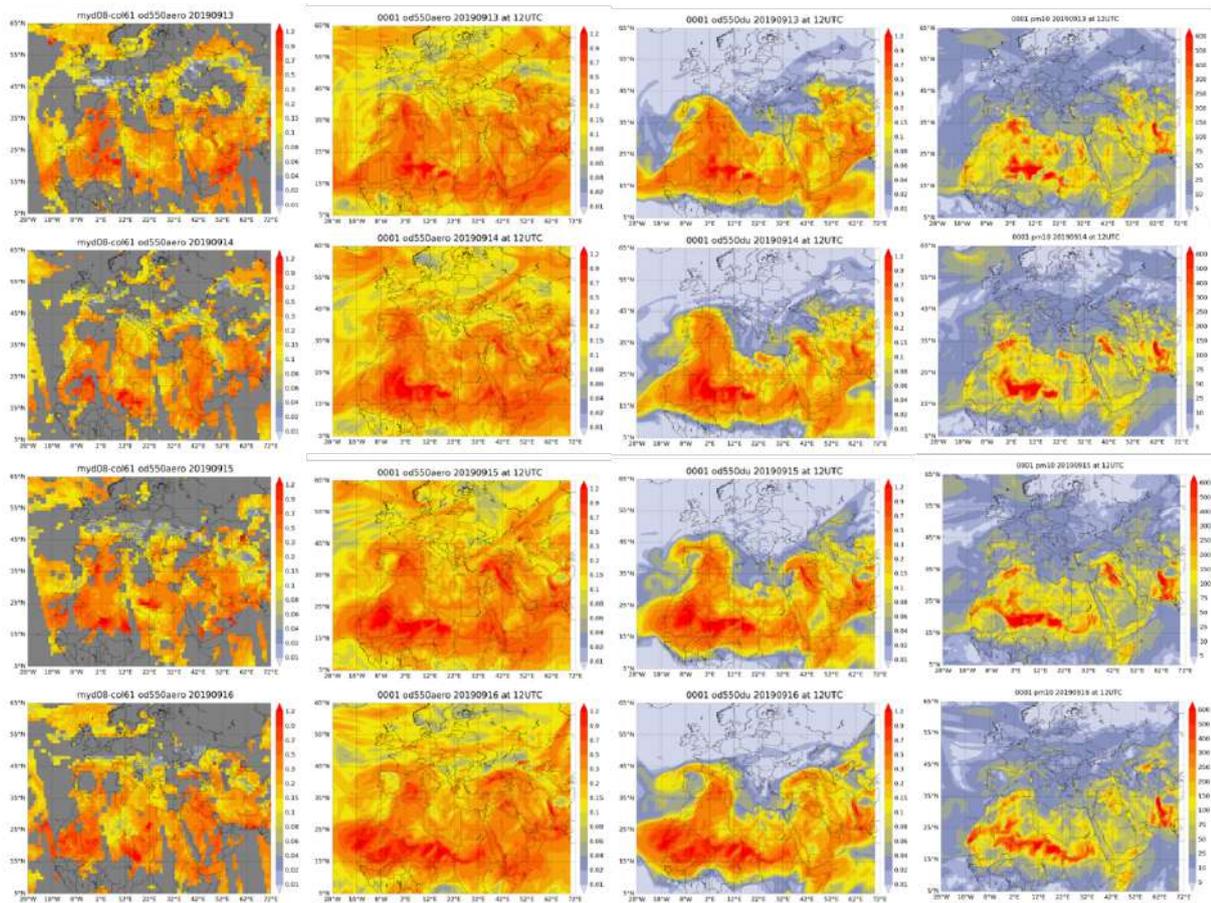
Figure 10.2.2. AOD from MODIS/Aqua Collection 6.1 Daily Level 3 product as well as AOD, DOD and PM10 at 12UTC from o-suite (central and left columns) for 12-16 September 2019.
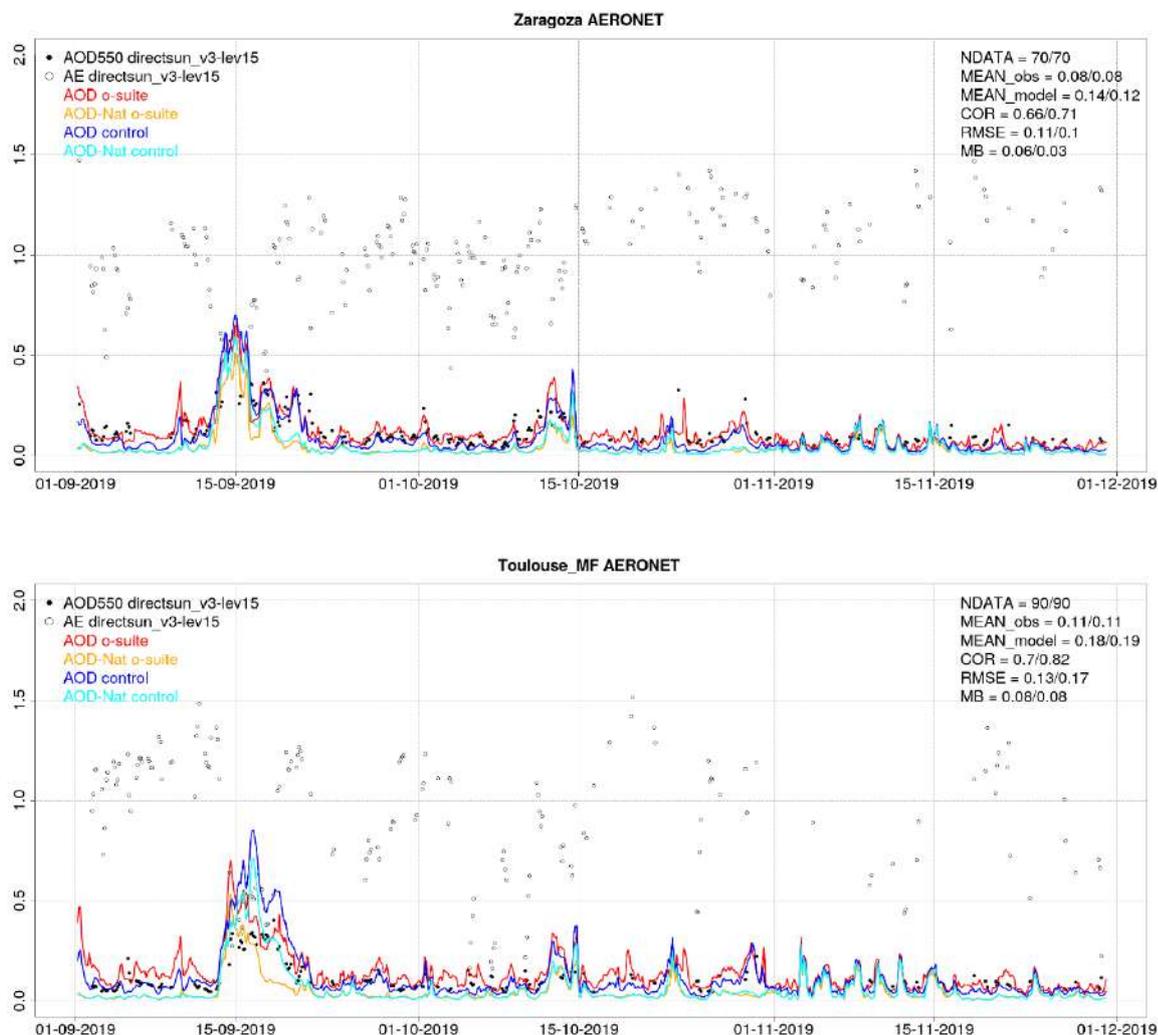
Figure 10.2.3: AOD from AERONET (black dot), AOD o-suite (red line), AOD control (blue line), AOD-Nat o-suite (orange line), AOD-Nat control (cyan line), for the study period Zaragoza (Central Iberian Peninsula) and Toulouse-MF (South France). AOD-Nat corresponds to the natural aerosol optical depth that includes dust and sea-salt. Skill scores per each individual site and model (o—suite/control) are shown in the upper right corner (NDATA: available 3-hourly values used for the calculations, MEAN observations, MEAN model, COR, RMSE, MB).

10.3 Dust event over Germany, 23 April 2019

After an exceptional drought in 2018, and a very dry vegetation season start in 2019, extreme drought conditions were measured in Poland and Germany. With a high-pressure system over Poland and large pressure gradients, south-westerly winds intensified and lead to local dust storms on the Baltic coast of Germany as reported in local newspaper and on TV. The dust was picked up
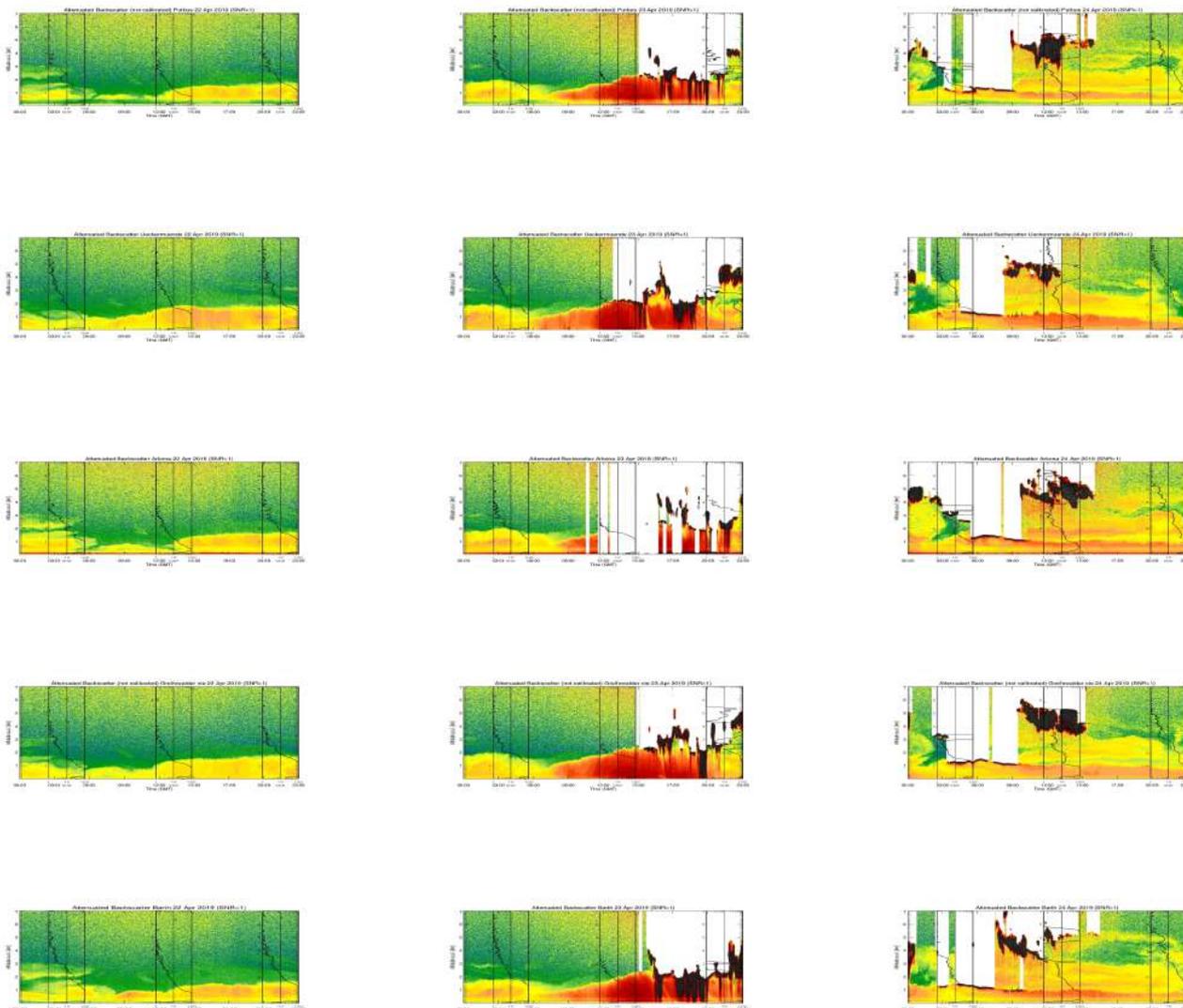
Figure 10.3.1. Sequence of the event at five Ceilometer stations for 22 April (left), 23 April (middle) and 24 April (right).

from dry soils in Poland and was transported along the pressure gradient. As similar case of Ukrainian Dust over central Europe was observed in 2007 (Birmli et al, 2008). By looking at the comparison of different dust model AODs, there is some indication for Saharan dust also transported in the area on the 24th and 25th of April. An extensive Satellite study was conducted by EUMETSAT, available here:

https://www.eumetsat.int/website/home/Images/ImageLibrary/DAT_4384629.html

Figure 10.3.1 shows the chronical sequence of the event of five different ceilometers in the area. Starting in the evening of the 22th, the event build up until reaching its maxima at 15 to 18 UTC on April the 23th. The event therefore lasts more than 24 hours. During that time the model show little guidance (Figs. 10.3.3. to 10.3.5). The main event between 9 and 21 UTC on 23th of April is
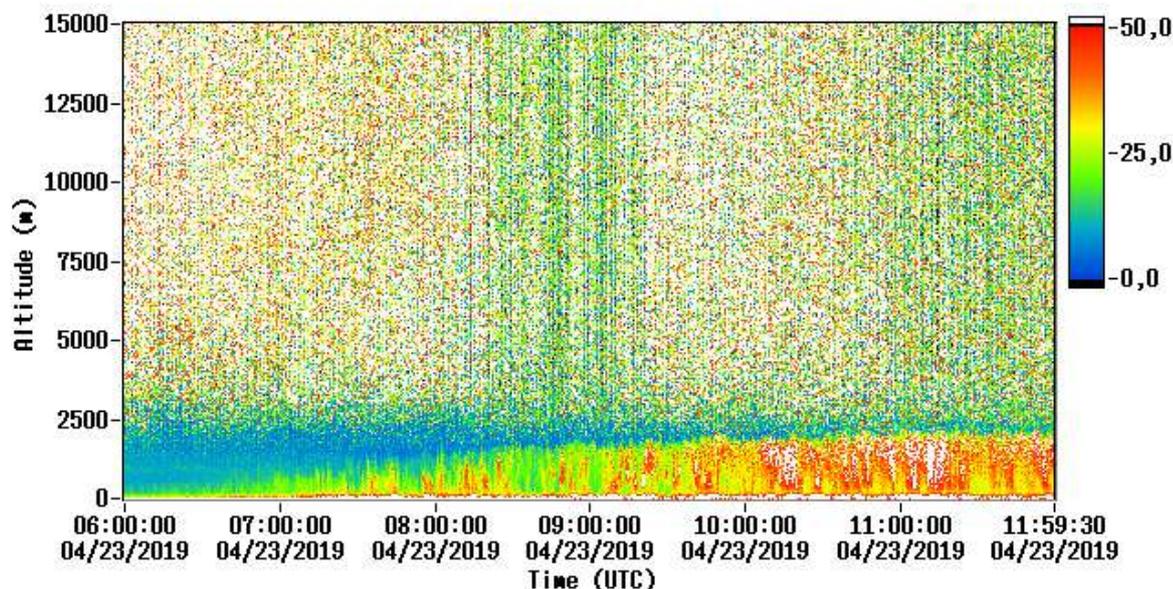
Figure 10.3.2. Lidar measurements from station Warsaw, on 23.4.2019

therefore not captured. This could mean that the model has little or no information about this event. Interestingly the model showed a response the following day (25.4), which is not fully supported by the ceilometer. The forward trajectories indicate that long range dust from the Saharan region may be transported into the area, but either arrived on 24th and 25th of April or were drifting over the UK, or were transported to heights well above 1500 m. The lidar depolarization ratio from Warsaw for the 23th of April shows that most of the depolarization material (e.g. dust and sand) is below 2500 m (Fig. 10.3.2), indicating that most of the material is transported in the boundary layer.

Backward trajectories for the air parcels originated at Putbus station indicates the transport from central/eastern Poland. Looking at satellite image NDVI, as indicator for the greenness of vegetation and therefore their stress, one can see that the eastern parts of Poland have very low NDVIs at the time of the event and therefore less active biomass, which makes soil more barren and prone for wind uplift. One can only speculate that the assimilation introduced some dust loading on the 24th, either from local sources or from long range transport. The difference in guidance between operational (red dots) and control run (without assimilation, blue dots) (see Fig. 10.3.5) indicates an introduction of aerosol mass during assimilation.

Conclusion:

Due to the prolonged dryness, vegetation activity in eastern Poland was very weak. Accompanied by wind gusts, dust has been uplifted and was transported along the pressure gradient in the boundary layer (up to 2500m). Shortly after (1 to 1,5 days) Saharan dust may have been transported to the area in the free troposphere and may have been sedimented due to rain events.
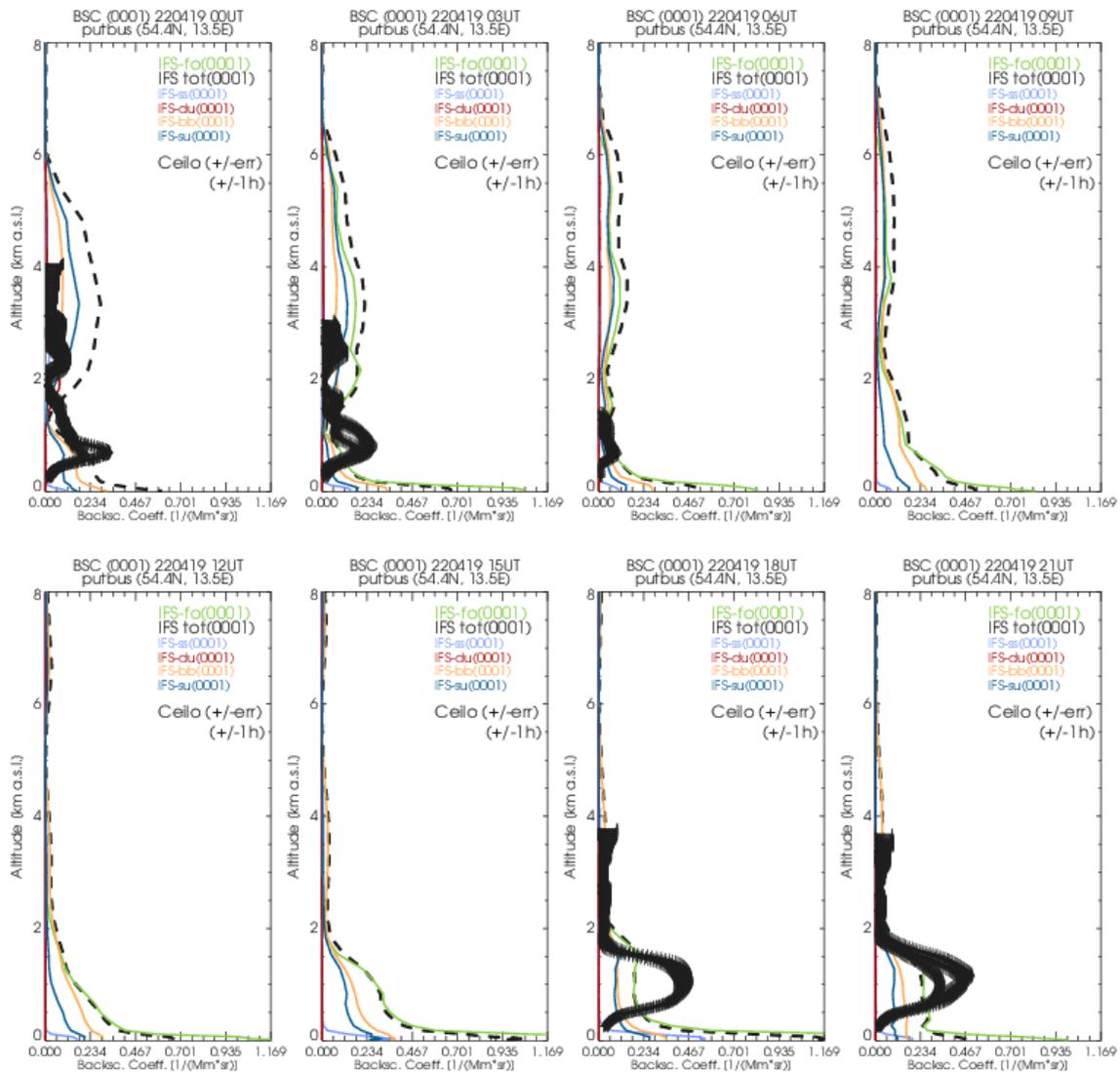
Figure 10.3.3. Backscatter coefficients for the operational run with assimilation, for 22.04.19.
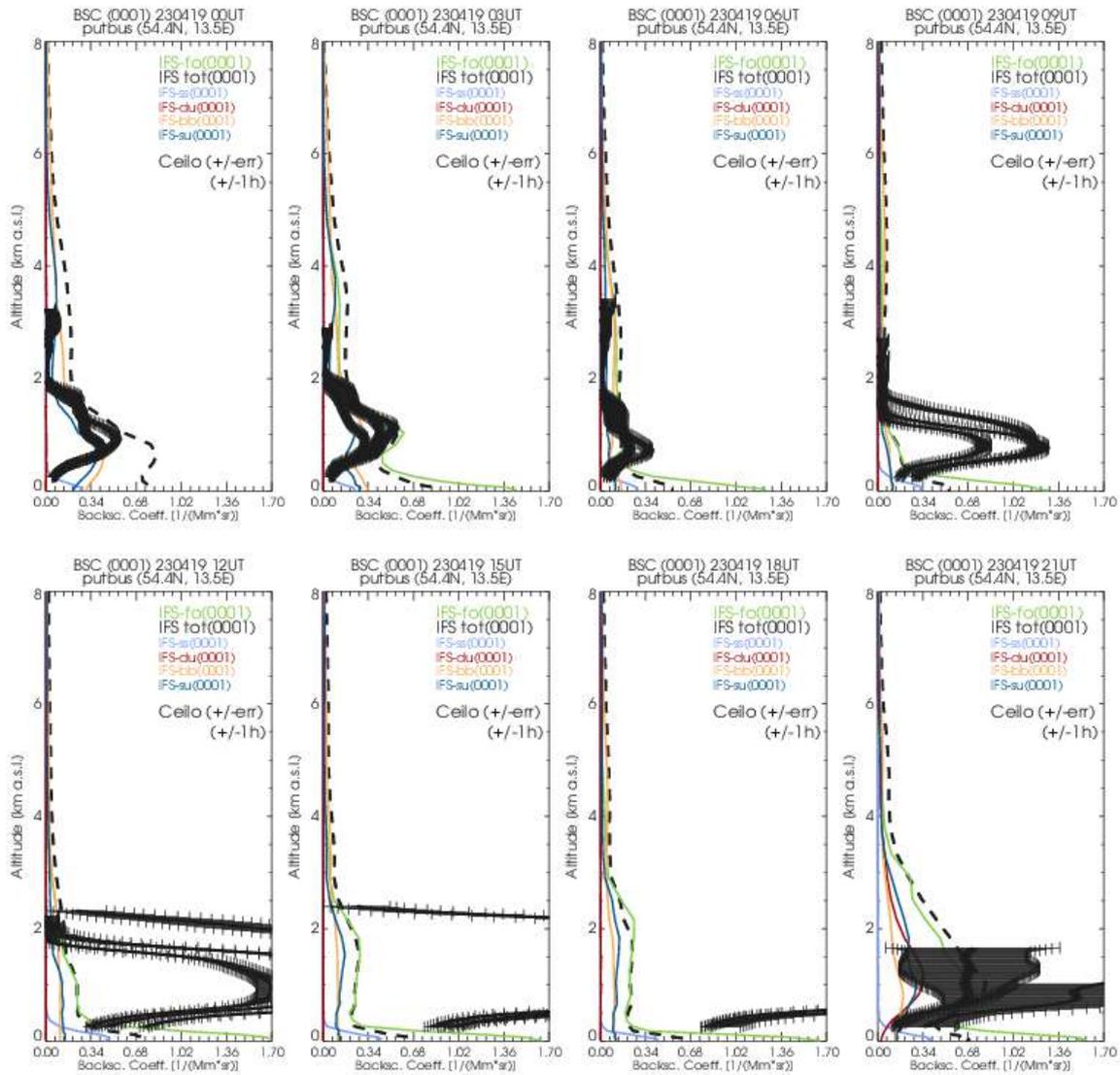
Figure 10.3.4. Backscatter coefficients for the operational run with assimilation, for 23.04.19.
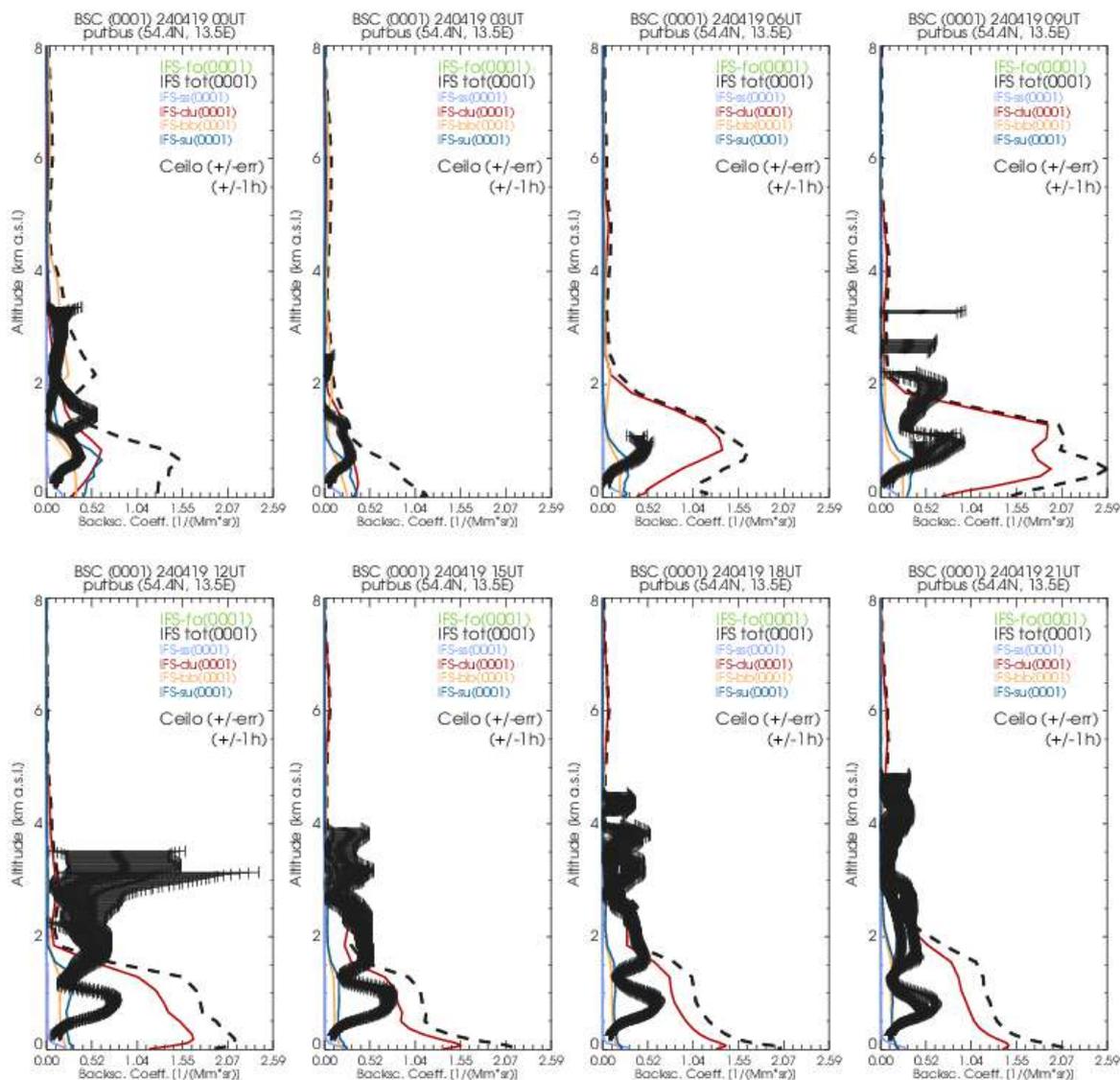
Figure 10.3.5. Backscatter coefficients for the operational run with assimilation, for 24.04.19.
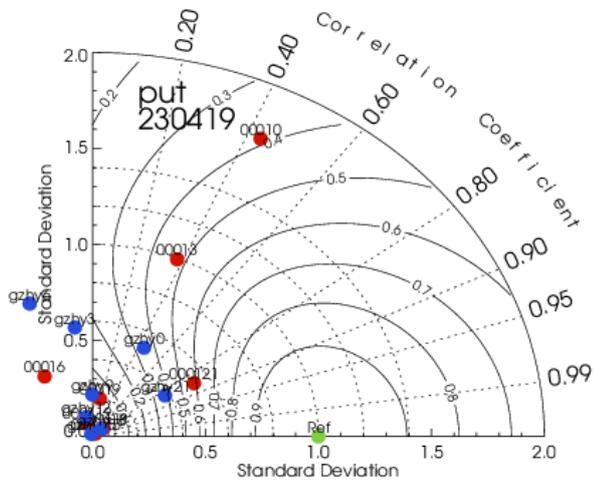
Figure 10.3.6. Taylorplot for Putbus for 23 April 2019.

11. References

Agusti-Panareda, A., Monitoring upgrades of analysis/forecast system, MACC-III Deliverable D44.04, June 2015.

Basart, S, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, E. Cuevas, H. J. Eskes, K. M. Hansen, O. Jorba, J. Kapsomenakis, B. Langerock, T. Pay, A. Richter, N. Sudarchikova, M. Schulz, A. Wagner, C. Zerefos, Upgrade verification note for the CAMS near-real time global atmospheric composition service: Evaluation of the e-suite for the CAMS upgrade of July 2019, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC1_D3.2.1-201907_esuite_v1.pdf, July 2019.

Bergamaschi, P., Frankenberg, C., Meirink, J. F., Krol, M., Villani, M. G., Houweling, S., Dentener, F., Dlugokencky, E. J., Miller, J. B., Gatti, L. V., Engel, A., and Levin, I.: Inverse modeling of global and regional CH4 emissions using SCIAMACHY satellite retrievals, J. Geophys. Res., 114, D22301, doi:10.1029/2009JD012287, 2009.

Benedetti, A., J.-J. Morcrette, O. Boucher, A. Dethof, R. J. Engelen, M. Fisher, H. Flentjes, N. Huneeus, L. Jones, J. W. Kaiser, S. Kinne, A. Mangold, M. Razinger, A. J. Simmons, M. Suttie, and the GEMS-AER team: Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part II : Data assimilation, J. Geophys. Res., 114, D13205, doi:10.1029/2008JD011115, 2009.

Birmili, W., Schepanski, K., Ansmann, A., Spindler, G., Tegen, I., Wehner, B., Nowak, A., Reimer, E., Mattis, I., Müller, K., Brüggemann, E., Gnauk, T., Herrmann, H., Wiedensohler, A., Althausen, D., Schladitz, A., Tuch, T., and Löschau, G.: A case of extreme particulate matter concentrations over Central Europe caused by dust emitted over the southern Ukraine, Atmos. Chem. Phys., 8, 997–1016, https://doi.org/10.5194/acp-8-997-2008, 2008.

Boussetta, S., Balsamo, G., Beljaars, A., Agusti-Panareda, A., Calvet, J.-C., Jacobs, C., van den Hurk, B., Viterbo, P., Lafont, S., Dutra, E., Jarlan, L., Balzarolo, M., Papale, D., and van der Werf, G.: Natural carbon dioxide exchanges in the ECMWF Integrated Forecasting System: implementation and offline validation, J. Geophys. Res.-Atmos., 118, 1–24, doi: 10.1002/jgrd.50488, 2013.

Braathen, WMO Arctic Ozone Bulletin No 1/2016, DOI:10.13140/RG.2.1.4929.6403, 2016.

Cammas, J.P., Brioude J., Chaboureau J.-P., Duron J., Mari C., Mascart P., Nédélec P., Smit H., Pätz H.-W., Volz-Thomas A., Stohl A., and Fromm M., Injection in the lower stratosphere of biomass fire emissions followed by long-range transport: a MOZAIC case study. Atmos. Chem. Phys., 9, 5829-5846, 2009

Cariolle, D. and Teyssèdre, H.: A revised linear ozone photochemistry parameterization for use in transport and general circulation models: multi-annual simulations, Atmos. Chem. Phys., 7, 2183-2196, doi:10.5194/acp-7-2183-2007, 2007.

Dee, D. P. and S. Uppala, Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. Quart. J. Roy. Meteor. Soc.,135, 1830-1841, 2009.

Deeter, M. N., Emmons, L. K., Edwards, D. P., Gille, J. C., and Drummond, J. R.: Vertical resolution and information content of CO profiles retrieved by MOPITT, Geophys. Res. Lett., 31, L15112, doi:10.1029/2004GL020235, 2004.

Deeter, M. N., et al. (2010), The MOPITT version 4 CO product: Algorithm enhancements, validation, and long-term stability, J. Geophys. Res., 115, D07306, doi:10.1029/2009JD013005.

Dentener, F., et al., 2006: Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom, Atmos. Chem. Phys., 6, 4321 – 4344.

Deshler, T., J.L. Mercer, H.G.J. Smit, R. Stubi, G. Levrat, B.J. Johnson, S.J. Oltmans, R. Kivi, A.M. Thompson, J. Witte, J. Davies, F.J. Schmidlin, G. Brothers, T. Sasaki (2008) Atmospheric comparison of electrochemical cell ozonesondes from different maufacturers, and with different cathode solution strengths: The Balloon Experiment on Standards for Ozonsondes. J. Geophys. Res.113, D04307, doi:10.1029/2007JD008975

Dupuy, E., et al.: Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE), Atmos. Chem. Phys., 9, 287-343, doi:10.5194/acp-9-287-2009, 2009.

Elbern, H., Schwinger, J., Botchorishvili, R.: Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: System configuration. Journal of Geophysical Research (Atmospheres) 115, 6302, 2010.

Emmons, L. K., D. P. Edwards, M. N. Deeter, J. C. Gille, T. Campos, P. Nédélec, P. Novelli, and G. Sachse, Measurements of Pollution In The Troposphere (MOPITT) validation through 2006 Atmos. Chem. Phys., 9, 1795-1803, 2009

Errera, Q., Daerden, F., Chabrillat, S., Lambert, J. C., Lahoz, W. A., Viscardy, S., Bonjean, S., and Fonteyn, D., 4D-Var Assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses, Atmos. Chem. Phys., 8, 6169-6187, 2008.

Errera, Q. and Ménard, R.: Technical Note: Spectral representation of spatial correlations in variational assimilation with grid point models and application to the belgian assimilation system for chemical observations (BASCOE), Atmos. Chem. Phys. Discuss., 12, 16763-16809, doi:10.5194/acpd-12-16763-2012, 2012.

Eskes, H.J., S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, E. Cuevas, H. Flentje, K. M. Hansen, J. Kapsomenakis, B. Langerock, M. Ramonet, A. Richter, M. Schulz, N. Sudarchikova, A. Wagner, T. Warneke, C. Zerefos, Observation characterisation and validation methods document, Copernicus Atmosphere Monitoring Service (CAMS) report, December 2019, doi: 10.24380/0nsd- wb26. Available from: http://atmosphere.copernicus.eu/user-support/validation/verification-global-services

Eskes, H. J., S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, Y. Christophe, H. Clark, E. Cuevas, K. M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, K. Petersen, M. Schulz, A. Wagner, C. Zerefos, Upgrade verification note for the CAMS near-real time global atmospheric composition service, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2015SC3_D84.3.1.5_201802_esuite_v1.pdf, February 2018 (2018b)

Eskes et al., Upgrade verification note for the CAMS near-real time global atmospheric composition service, Addendum July 2018, CAMS84_2015SC3_D84.3.1.5_201802_esuite_v1.pdf (2018c).

Flemming, J., Huijnen, V., Arteta, J., Bechtold, P., Beljaars, A., Blechschmidt, A.-M., Diamantakis, M., Engelen, R. J., Gaudel, A., Inness, A., Jones, L., Josse, B., Katragkou, E., Marecal, V., Peuch, V.-H., Richter, A., Schultz, M. G., Stein, O., and Tsikerdekis, A.: Tropospheric chemistry in the Integrated Forecasting System of ECMWF, Geosci. Model Dev., 8, 975-1003, doi:10.5194/gmd-8-975-2015, 2015.

Flemming, J., Benedetti, A., Inness, A., Engelen, R. J., Jones, L., Huijnen, V., Remy, S., Parrington, M., Suttie, M., Bozzo, A., Peuch, V.-H., Akritidis, D., and Katragkou, E.: The CAMS interim Reanalysis of Carbon Monoxide, Ozone and Aerosol for 2003–2015, Atmos. Chem. Phys., 17, 1945-1983, doi:10.5194/acp-17-1945-2017, 2017.

Franco, B., et al., Retrievals of formaldehyde from ground-based FTIR and MAX-DOAS observations at the Jungfraujoch station and comparisons with GEOS-Chem and IMAGES model simulations, Atmos. Meas. Tech., 8, 1733-1756, 2015

Gielen, C., Van Roozendael, M., Hendrick, F., Pinardi, G., Vlemmix, T., De Bock, V., De Backer, H., Fayt, C., Hermans, C., Gillotay, D., and Wang, P.: A simple and versatile cloud-screening method for MAX-DOAS retrievals, Atmos. Meas. Tech., 7, 3509-3527, doi:10.5194/amt-7-3509-2014, 2014.

Granier, C. et al.: Evolution of anthropogenic and biomass burning emissions of air pollutants at global and regional scales during the 1980–2010 period. Climatic Change (109), 2011

Holben, B. N., Eck, T. F., Slutsker, I., Tanré, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov A.: AERONET – a federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1–16, 5529, 5533, 5537, 5544, 1998.

Hommel, R., Eichmann, K.-U., Aschmann, J., Bramstedt, K., Weber, M., von Savigny, C., Richter, A., Rozanov, A., Wittrock, F., Khosrawi, F., Bauer, R., and Burrows, J. P.: Chemical ozone loss and ozone mini-hole event during the Arctic winter 2010/2011 as observed by SCIAMACHY and GOME-2, Atmos. Chem. Phys., 14, 3247-3276, doi:10.5194/acp-14-3247-2014, 2014.

Huijnen, V., et al.: The global chemistry transport model TM5: description and evaluation of the tropospheric chemistry version 3.0, Geosci. Model Dev., 3, 445-473, doi:10.5194/gmd-3-445-2010, 2010.

Inness, A., Blechschmidt, A.-M., Bouarar, I., Chabrillat, S., Crepulja, M., Engelen, R. J., Eskes, H., Flemming, J., Gaudel, A., Hendrick, F., Huijnen, V., Jones, L., Kapsomenakis, J., Katragkou, E., Keppens, A., Langerock, B., de Mazière, M., Melas, D., Parrington, M., Peuch, V. H., Razinger, M., Richter, A., Schultz, M. G., Suttie, M., Thouret, V., Vrekoussis, M., Wagner, A., and Zerefos, C.: Data assimilation of satellite-retrieved ozone, carbon monoxide and nitrogen dioxide with ECMWF's Composition-IFS, Atmos. Chem. Phys., 15, 5275-5303, doi:10.5194/acp-15-5275-2015, 2015.

Janssens-Maenhout, G., Dentener, F., Aardenne, J. V., Monni, S., Pagliari, V., Orlandini, L., Klimont, Z., Kurokawa, J., Akimoto, H., Ohara, T., Wankmueller, R., Battye, B., Grano, D., Zuber, A., and Keating, T.: EDGAR-HTAP: a Harmonized Gridded Air Pollution Emission Dataset Based on National Inventories, JRC68434, EUR report No EUR 25 299–2012, ISBN 978-92-79- 23122-0, ISSN 1831-9424, European Commission Publications Office, Ispra (Italy), 2012.

Jaross, G., Bhartia, P.K., Chen, G., Kowitt, M., Haken, M., Chen, Z., Xu, Ph., Warner, J., Kelly, T. : OMPS Limb Profiler instrument performance assessment, J. Geophys. Res. Atmos 119, 2169-8996, 2014.

Kaiser, J. W., Heil, A., Andreae, M. O., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M. G., Suttie, M., and van der Werf, G. R.: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power, Biogeosciences, 9, 527-554, doi:10.5194/bg-9-527-2012, 2012.

Kramarova, N. A., Nash, E. R., Newman, P. A., Bhartia, P. K., McPeters, R. D., Rault, D. F., Seftor, C. J., Xu, P. Q., and Labow, G. J.: Measuring the Antarctic ozone hole with the new Ozone Mapping and Profiler Suite (OMPS), Atmos. Chem. Phys., 14, 2353-2361, doi:10.5194/acp-14-2353-2014, 2014.

Lahoz, W. A., Errera, Q., Viscardy, S., and Manney G. L., The 2009 stratospheric major warming described from synergistic use of BASCOE water vapour analyses and MLS observations, Atmos. Chem. Phys. 11, 4689-4703, 2011

Lambert, A, et al., Aura Microwave Limb Sounder Version 3.4 Level-2 near real-time data user guide, http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/MLS/documents/NRT-user-guide-v34.pdf

Langerock, B., De Mazière, M., Hendrick, F., Vigouroux, C., Desmet, F., Dils, B., and Niemeijer, S.: Description of algorithms for co-locating and comparing gridded model data with remote-sensing observations, Geosci. Model Dev., 8, 911-921, doi:10.5194/gmd-8-911-2015, 2015.

Lefever, K., van der A, R., Baier, F., Christophe, Y., Errera, Q., Eskes, H., Flemming, J., Inness, A., Jones, L., Lambert, J.-C., Langerock, B., Schultz, M. G., Stein, O., Wagner, A., and Chabrillat, S.: Copernicus stratospheric ozone service, 2009–2012: validation, system intercomparison and roles of input data sets, Atmos. Chem. Phys., 15, 2269-2293, doi:10.5194/acp-15-2269-2015, 2015.

Liu, Z., et al., Exploring the missing source of glyoxal (CHOCHO) over China, Geophys. Res. Lett., 39, L10812, doi: 10.1029/2012GL051645, 2012

Massart, S., Flemming, J., Cariolle, D., Jones, L., High resolution CO tracer forecasts, MACC-III Deliverable D22.04, May 2015, available from http://www.gmes-atmosphere.eu/documents/macciii/deliverables/grg

Morcrette, J.-J., O. Boucher, L. Jones, D. Salmond, P. Bechtold, A. Beljaars, A. Benedetti, A. Bonet, J. W. Kaiser, M. Razinger, M. Schulz, S. Serrar, A. J. Simmons, M. Sofiev, M. Suttie, A. M. Tompkins, and A. Untch: Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part I: Forward modelling, J. Geophys. Res., 114, D06206, doi:10.1029/2008JD011235, 2009.

Ramonet, M., A. Wagner, M. Schulz, Y. Christophe, H. J. Eskes, S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, E. Cuevas, A. El-Yazidi, H. Flentje, K.M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, A. Richter, N. Sudarchikova, V. Thouret, T. Warneke, C. Zerefos, Validation report of the CAMS near-real-time global atmospheric composition service: Period June - August 2019, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC1_D1.1.1_JJA2019_v1.pdf, November 2019.

Rémy, S., Kipling, Z., Flemming, J., Boucher, O., Nabat, P., Michou, M., Bozzo, A., Ades, M., Huijnen, V., Benedetti, A., Engelen, R., Peuch, V.-H., and Morcrette, J.-J.: Description and evaluation of the tropospheric aerosol scheme in the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS-AER, cycle 45R1), Geosci. Model Dev., 12, 4627–4659, https://doi.org/10.5194/gmd-12-4627-2019, 2019.

Richter, A., Burrows, J. P., Nüß, H., Granier, C, Niemeier, U.: Increase in tropospheric nitrogen dioxide over China observed from space, Nature, 437, 129-132, doi: 10.1038/nature04092, 2005

Richter, A., Begoin, M., Hilboll, A., and Burrows, J. P.: An improved NO2 retrieval for the GOME-2 satellite instrument, Atmos. Meas. Tech. , **4**, 1147-1159, doi:10.5194/amt-4-1147-2011, 2011

Schulz, M., Y. Christophe, M. Ramonet, Wagner, A., H.J. Eskes, S. Basart, A. Benedictow, Y. Bennouna, A.-M. Blechschmidt, S. Chabrillat, E. Cuevas, A. El- Yazidi, H. Flentje, K.M. Hansen, U. Im, J. Kapsomenakis, B. Langerock, A. Richter, N. Sudarchikova, V. Thouret, T. Warneke, C. Zerefos, Validation report of the CAMS near-real-time global atmospheric composition service: Period December 2018 -February 2019, Copernicus Atmosphere Monitoring Service (CAMS) report, CAMS84_2018SC1_D1.1.1_DJF2019_v1.pdf, June 2019, doi:10.24380/7th6-tk72.

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., Müller, J.-F., Kuhn, U., Stefani, P., and Knorr, W.: Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, Atmos. Chem. Phys., 14, 9317-9341, doi:10.5194/acp-14-9317-2014, 2014.

Smit, H.G.J., W. Straeter, B.J. Johnson, S.J. Oltmans, J. Davies, D.W. Tarasick, B. Hoegger, R. Stubi, F.J. Schmidlin, T. Northam, A.M. Thompson, J.C. Witte, I. Boyd: Assessment of the performance of ECC-ozonesondes under quasi-flight conditions in the environmental simulation chamber: Insights from the Juelich Ozone Sonde Intercomparison Experiment (JOSIE), J. Geophys. Res. 112, D19306, doi:10.1029/2006JD007308, 2007.

Solomon, S., Haskins, J., Ivy, D. J. and Min, F.: Fundamental differences between Arctic and Antarctic ozone depletion, PNAS 2014 111 (17) 6220-6225, doi:10.1073/pnas.1319307111, 2014.

Stavrakou,T., First space-based derivation of the global atmospheric methanol fluxes, Atm. Chem. Phys., 11, 4873-4898, 2013.

Strahan, S.E., A.R. Douglass, and P.A. Newman, The contributions of chemistry and transport to low arctic ozone in March 2011 derived from Aura MLS observations, J. Geophys. Res. Atmos., 118, 1563–1576, doi:10.1002/jgrd.50181, 2013.

Taha, G.; Jaross, G. R.; Bhartia, P. K.: Validation of OMPS LP Ozone Profiles Version 2.0 with MLS, Ozone Sondes and Lidar Measurements, American Geophysical Union, Fall Meeting 2014, abstract #A33J-3322, 2014.

Taylor, K.E.: Summarizing multiple aspects of model performance in a single diagram. J.Geophys. Res., 106, 7183-7192, 2001.

van der A, R. J. , M. A. F. Allaart, and H. J. Eskes, Multi sensor reanalysis of total ozone, Atmos. Chem. Phys., 10, 11277–11294, doi:10.5194/acp-10-11277-2010, www.atmos- chem-phys.net/10/11277/2010/, 2010

van der A, R., M. Allaart, H. Eskes, K. Lefever, Validation report of the MACC 30-year multi-sensor reanalysis of ozone columns Period 1979-2008, MACC-II report, Jan 2013, MACCII_VAL_DEL_D_83.3_OzoneMSRv1_20130130.docx/pdf.

van der A, R. J., Allaart, M. A. F., and Eskes, H. J.: Extended and refined multi sensor reanalysis of total ozone for the period 1970–2012, Atmos. Meas. Tech., 8, 3021-3035, doi:10.5194/amt-8-3021-2015, 2015.

*Vrekoussis, M., Wittrock, F., Richter, A., and Burrows, J. P.: GOME-2 observations of oxygenated VOCs: what can we learn from the ratio glyoxal to formaldehyde on a global scale?, Atmos. Chem. Phys., **10**, 10145-10160, doi:10.5194/acp-10-10145-2010, 2010*

Wennberg, P. O., Mui, W., Wunch, D., Kort, E. A., Blake, D. R., Atlas, E. L., Santoni, G. W., Wofsy, S. C., Diskin, G. S., Jeong, S., and Fischer, M. L.: On the sources of methane to the Los Angeles atmosphere, Environ. Sci. Technol., 46, 9282–9289, https://doi.org/10.1021/es301138y, 2012

Wittrock, F., A. Richter, H. Oetjen, J. P. Burrows, M. Kanakidou, S. Myriokefalitakis, R. Volkamer, S. Beirle, U. Platt, and T. Wagner, Simultaneous global observations of glyoxal and formaldehyde from space, Geophys. Res. Lett., 33, L16804, doi:10.1029/2006GL026310, 2006

WMO (2010), Guidelines for the Measurement of Atmospheric Carbon Monoxide, GAW Report No. 192, World Meteorological Organization, Geneva, Switzerland, 2010.

WMO (2013), Guidelines for the Continuous Measurements of Ozone in the Troposphere, GAW Report No. 209, World Meteorological Organization, Geneva, Switzerland, 2013.

Wunch, D., Wennberg, P. O., Toon, G. C., Keppel-Aleks, G., and Yavin, Y. G.: Emissions of greenhouse gases from a North American megacity, Geophys. Res. Lett., 36, 1–5, https://doi.org/10.1029/2009GL039825, 2009.

Annex 1: Acknowledgements

Listed below are the authors contributing to the sections in this report. The authors contributing to the model description are also provided, as well as acknowledgements to the validation datasets.

Tropospheric reactive gases reactive gases

Annette Wagner, MPG (editor, O3 sondes, GAW data)
Yasmine Bennouna, Valerie Thouret, CNRS-LA (IAGOS)
Harald Flentje, DWD (O3 sondes, GAW data)
Anne Blechschmidt and Andreas Richter, IUB Bremen (GOME-2 NO2, HCHO)
John Kapsomenakis, Christos Zerefos, AA (ESRL)
Natalia Sudarchikova, satellite IR observations (MPG)
Kaj Hansen, Ulas Im, AU (Arctic theme)
Bavo Langerock, BIRA (NDACC)

Tropospheric aerosol

Michael Schulz, MetNo (editor, Aerocom, Aeronet)
Anna Benedictow, Jan Griesfeller, MetNo (Aerocom, Aeronet)
Sara Basart, MTeresa Pay, Oriol Jorba, BSC-CNS (Aeronet, MODIS, AirBase, SDS-WAS NAMEE RC)
Emilio Cuevas, AEMET (Aeronet, MODIS, AirBase, SDS-WAS NAMEE RC)
Harald Flentje, DWD (Backscatter profiles)

Stratospheric reactive gases

Yves Christophe, BIRA (editor, model-satellite intercomparisons)
Simon Chabrillat, BIRA (model intercomparisons)
Annette Wagner, MPI-M (O$_3$ sondes)
Bavo Langerock, BIRA (NDACC FTIR, MWR, UVVIS DOAS, LIDAR)
Anne Blechschmidt and Andreas Richter, IUB-UB Bremen (SCIAMACHY/GOME-2 NO$_2$)

Greenhouse gases

Michel Ramonet, IPSL (ICOS)
Abdelhadi El-Yazidi and Leonard Rivier, LSCE (ICOS)
Thorsten Warneke, UBC (TCCON)
Bavo Langerock, BIRA (TCCON)

Reactive gases and aerosol modeling

Johannes Flemming (ECMWF), Antje Inness (ECMWF), Angela Benedetti (ECMWF), Sebastien Massart (ECMWF), Anna Agusti-Panareda (ECMWF), Johannes Kaiser (KCL/MPIC/ECMWF), Samuel Remy (LMD), Olivier Boucher (LMD), Vincent Huijnen (KNMI), Richard Engelen (ECMWF)

Acknowledgements for the validation datasets used

TCCON references:

Hazan, L., J. Tarniewicz, M. Ramonet, O. Laurent and A. Abbaris **(2016)**. *Automatic processing of atmospheric CO2 and CH4 mole fractions at the ICOS Atmosphere Thematic Centre*. Atmospheric Measurement Techniques 9(9): 4719-4736.

Blumenstock, T., F. Hase, M. Schneider, O. E. García, and E. Sepúlveda. 2017. "TCCON data from Izana (ES), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.izana01.r1.

De Mazière, M., M. K. Sha, F. Desmet, C. Hermans, F. Scolas, N. Kumps, J.-M. Metzger, V. Duflot, and J.-P. Cammas. 2017. "TCCON data from Réunion Island (RE), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.reunion01.r1.

Deutscher, N. M., J. Notholt, J. Messerschmidt, C. Weinzierl, T. Warneke, C. Petri, and P. Grupe. 2017. "TCCON data from Bialystok (PL), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.bialystok01.r1/1183984.

Dubey, M. K., B. G. Henderson, D. Green, Z. T. Butterfield, G. Keppel-Aleks, N. T. Allen, J.-F. Blavier, C. M. Roehl, D. Wunch, and R. Lindenmaier. 2017. "TCCON data from Manaus (BR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.manaus01.r0/1149274.

Dubey, M. K., R. Lindenmaier, B. G. Henderson, D. Green, N. T. Allen, C. M. Roehl, J.-F. Blavier, et al. 2017. "TCCON data from Four Corners (US), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.fourcorners01.r0/1149272.

Feist, D. G., S. G. Arnold, N. John, and M. C. Geibel. 2017. "TCCON data from Ascension Island (SH), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.ascension01.r0/1149285.

Goo, T.-Y., Y.-S. Oh, and V. A. Velazco. 2017. "TCCON data from Anmeyondo (KR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.anmeyondo01.r0/1149284.

Griffith, D. W. T., N. M. Deutscher, V. A. Velazco, P. O. Wennberg, Y. Yavin, G. Keppel-Aleks, R. A. Washenfelder, et al. 2017. "TCCON data from Darwin (AU), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.darwin01.r0/1149290.

Griffith, D. W. T., V. A. Velazco, N. M. Deutscher, C. Paton-Walsh, N. B. Jones, S. R. Wilson, R. C. Macatangay, G. C. Kettlewell, R. R. Buchholz, and M. O. Riggenbach. 2017. "TCCON data from Wollongong (AU), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.wollongong01.r0/1149291.

Hase, F., T. Blumenstock, S. Dohe, J. Groß, and M.ä. Kiel. 2017. "TCCON data from Karlsruhe (DE), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.karlsruhe01.r1/1182416.

Iraci, L. T., J. R. Podolske, P. W. Hillyard, C. Roehl, P. O. Wennberg, J.-F. Blavier, J. Landeros, et al. 2017. "TCCON data from Edwards (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.edwards01.r1/1255068.

—. 2017. "TCCON data from Indianapolis (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.indianapolis01.r1/1330094.

Kawakami, S., H. Ohyama, K. Arai, H. Okumura, C. Taura, T. Fukamachi, and M. Sakashita. 2017. "TCCON data from Saga (JP), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.saga01.r0/1149283.

Kivi, R., P. Heikkinen, and E. Kyrö. 2017. "TCCON data from Sodankylä (FI), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.sodankyla01.r0/1149280.

Liu, Cheng, Wei Wang, and Youwen Sun. 2018. "TCCON data from Hefei (PRC), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.hefei01.r0.

Morino, I., N. Yokozeki, T. Matsuzaki, and M. Horikawa. 2017. "TCCON data from Rikubetsu (JP), Release GGG2014.R2." CaltechDATA. doi:10.14291/tccon.ggg2014.rikubetsu01.r2.

Morino, I., T. Matsuzaki, and M. Horikawa. 2017. "TCCON data from Tsukuba (JP), 125HR, Release GGG2014.R2." CaltechDATA. doi:10.14291/tccon.ggg2014.tsukuba02.r2.

Morino, Isamu, Voltaire A. Velazco, Akihiro Hori, Osamu Uchino, and David W. T. Griffith. 2018. "TCCON data from Burgos, Ilocos Norte (PH), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.burgos01.r0.

Notholt, J., C. Petri, T. Warneke, N. M. Deutscher, M. Palm, M. Buschmann, C. Weinzierl, R. C. Macatangay, and P. Grupe. 2017. "TCCON data from Bremen (DE), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.bremen01.r0/1149275.

Notholt, J., T. Warneke, C. Petri, N. M. Deutscher, C. Weinzierl, M. Palm, and M. Buschmann. 2017. "TCCON data from Ny Ålesund, Spitsbergen (NO), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.nyalesund01.r0/1149278.

Pollard, David Frank, John Robinson, and Hisako Shiona. 2019. "TCCON data from Lauder (NZ), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.lauder03.r0.

Sherlock, V., B. Connor, J. Robinson, H. Shiona, D. Smale, and D. F. Pollard. 2017. "TCCON data from Lauder (NZ), 120HR, Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.lauder01.r0/1149293.

—. 2017. "TCCON data from Lauder (NZ), 125HR, Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.lauder02.r0/1149298.

Strong, K., S. Roche, J. E. Franklin, J. Mendonca, E. Lutsch, D. Weaver, P. F. Fogal, J. R. Drummond, R. Batchelor, and R. Lindenmaier. 2018. "TCCON data from Eureka (CA), Release GGG2014.R3." CaltechDATA. doi:10.14291/tccon.ggg2014.eureka01.r3.

Sussmann, R., and M. Rettinger. 2017. "TCCON data from Garmisch (DE), Release GGG2014.R2." CaltechDATA. doi:10.14291/tccon.ggg2014.garmisch01.r2.

—. 2018. "TCCON data from Zugspitze (DE), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.zugspitze01.r1.

Té, Y., P. Jeseck, and C. Janssen. 2017. "TCCON data from Paris (FR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.paris01.r0/1149279.

Warneke, T., J. Messerschmidt, J. Notholt, C. Weinzierl, N. M. Deutscher, C. Petri, and P. Grupe. 2017. "TCCON data from Orléans (FR), Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.orleans01.r0/1149276.

Wennberg, P. O., C. M. Roehl, D. Wunch, G. C. Toon, J.-F. Blavier, R. Washenfelder, G. Keppel-Aleks, N. T. Allen, and J. Ayers. 2017. "TCCON data from Park Falls (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.parkfalls01.r1.

Wennberg, P. O., C. M. Roehl, J.-F. Blavier, D. Wunch, and N. T. Allen. 2017. "TCCON data from Jet Propulsion Laboratory (US), 2011, Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.jpl02.r1/1330096.

Wennberg, P. O., D. Wunch, C. M. Roehl, J.-F. Blavier, G. C. Toon, and N. T. Allen. 2017. "TCCON data from Caltech (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.pasadena01.r1/1182415.

—. 2017. "TCCON data from Lamont (US), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.lamont01.r1/1255070.

Wennberg, P. O., D. Wunch, Y. Yavin, G. C. Toon, J.-F. Blavier, N. T. Allen, and G. Keppel-Aleks. 2017. "TCCON data from Jet Propulsion Laboratory (US), 2007, Release GGG2014.R0." CaltechDATA. doi:10.14291/tccon.ggg2014.jpl01.r0/1149163.

Wunch, D., J. Mendonca, O. Colebatch, N. T. Allen, J.-F. Blavier, S. Roche, J. Hedelius, et al. 2017. "TCCON data from East Trout Lake, SK (CA), Release GGG2014.R1." CaltechDATA. doi:10.14291/tccon.ggg2014.easttroutlake01.r1.

Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, C., Feist, D. G., & Wennberg, P. O. (2015). The Total Carbon Column Observing Network's GGG2014 Data Version. Tech. rep., California Institute of Technology, Pasadena. doi:10.14291/tccon.ggg2014.documentation.R0/1221662

Copernicus Atmosphere Monitoring Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-atmosphere.eu